

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, blastn frame, p2n, p2n, mode

Run on: April 6, 2006, 09:51:12 ; Search time 6373 Seconds

(without alignments)
3380.459 Million cell updates/sec

Title: US-10-780-703-2

Perfect score: 2034

Sequence: 1 MAIVSSTTSIIIPMSQVNNN.....MNPVDPGNDRHYTNIPFK 379

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame, p2n, model -DEV=xlh
-Q=/abse/ABSWEB_spool/US10780703/runat_06042006_101337_6322/app_query.fasta.1
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs03h
-USER=US10780703 @CGN_1_1_4939@runat_06042006_101337_6322 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*

1: gb_ba:*

2: gb_in:*

3: gb_env:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pr:*

9: gb_ro:*

10: gb_scs:*

11: gb_sy:*

12: gb_un:*

13: gb_vl:*

14: gb_htg:*

15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Length | ID | Description |
|------------|-------|--------------|----------|--------------------|
| 1 | 2034 | 100.0 | 1171 15 | BT000874 Arabidops |
| 2 | 1735 | 85.3 | 74903 15 | AC005312 Arabidops |
| 3 | 1574 | 77.4 | 1062 15 | AF360308 Arabidops |

| | | | | | | |
|----|-------|------|--------|----|----------|--------------------|
| 4 | 815.5 | 40.1 | 1631 | 15 | AK102902 | AK102902 Oryza sat |
| 5 | 796.5 | 39.2 | 1640 | 15 | AK072275 | AK072275 Oryza sat |
| 6 | 796.5 | 39.2 | 1891 | 15 | AK071052 | AK071052 Oryza sat |
| 7 | 789.5 | 38.8 | 2424 | 15 | AK068393 | AK068393 Oryza sat |
| 8 | 783 | 38.5 | 1406 | 15 | AY672069 | AY672069 Hordeum v |
| 9 | 781 | 38.4 | 1212 | 6 | AX654050 | AX654050 Sequence |
| 10 | 769.5 | 37.8 | 80577 | 14 | AP007649 | AP007649 Lotus cor |
| 11 | 769.5 | 37.8 | 91718 | 14 | AP007409 | AP007409 Lotus cor |
| 12 | 759 | 37.3 | 1250 | 6 | AX654570 | AX654570 Sequence |
| 13 | 688.5 | 33.8 | 549 | 6 | AX653764 | AX653764 Sequence |
| 14 | 672.5 | 33.1 | 110000 | 15 | AP008207 | Continuation (386 |
| 15 | 672.5 | 33.1 | 110000 | 15 | AP008207 | Continuation (387 |
| 16 | 672.5 | 33.1 | 150064 | 15 | AP003346 | AP003346 Oryza sat |
| 17 | 672.5 | 33.1 | 191022 | 15 | AP003431 | AP003431 Oryza sat |
| 18 | 655.5 | 32.2 | 110000 | 15 | AP008211 | Continuation (204 |
| 19 | 655.5 | 32.2 | 173074 | 15 | AC137623 | AC137623 Oryza sat |
| 20 | 611 | 30.0 | 110000 | 15 | AP008214 | Continuation (8 of |
| 21 | 611 | 30.0 | 185545 | 15 | AP005657 | AP005657 Oryza sat |
| 22 | 550 | 27.0 | 1271 | 15 | AK107090 | AK107090 Oryza sat |
| 23 | 546.5 | 26.9 | 1591 | 15 | AK073013 | AK073013 Oryza sat |
| 24 | 544.5 | 26.8 | 1158 | 6 | AX654270 | AX654270 Sequence |
| 25 | 536 | 26.4 | 950 | 6 | AX653803 | AX653803 Sequence |
| 26 | 517.5 | 25.4 | 1152 | 6 | AX652795 | AX652795 Sequence |
| 27 | 517 | 25.4 | 1423 | 15 | AY573802 | AY573802 Lycopersi |
| 28 | 516 | 25.4 | 1152 | 15 | DQ028772 | DQ028772 Glycine m |
| 29 | 515 | 25.3 | 978 | 15 | AF361804 | AF361804 Arabidops |
| 30 | 514 | 25.3 | 830 | 6 | AX651430 | AX651430 Sequence |
| 31 | 514 | 25.3 | 1021 | 15 | DQ028774 | DQ028774 Glycine m |
| 32 | 512 | 25.2 | 1567 | 15 | AF011555 | AF011555 Lycopersi |
| 33 | 512 | 25.2 | 1664 | 15 | AJ704829 | AJ704829 Prunus pe |
| 34 | 509 | 25.0 | 1209 | 6 | AR569436 | AR569436 Sequence |
| 35 | 509 | 25.0 | 1209 | 6 | AX087867 | AX087867 Sequence |
| 36 | 509 | 25.0 | 1565 | 15 | AY704177 | AY704177 Solanum t |
| 37 | 508.5 | 25.0 | 972 | 15 | BT008346 | BT008346 Arabidops |
| 38 | 508.5 | 25.0 | 1125 | 6 | CS137954 | CS137954 Sequence |
| 39 | 508.5 | 25.0 | 1336 | 15 | AK117878 | AK117878 Arabidops |
| 40 | 508 | 25.0 | 894 | 6 | CS137686 | CS137686 Sequence |
| 41 | 508 | 25.0 | 925 | 15 | AY091428 | AY091428 Arabidops |
| 42 | 508 | 25.0 | 1150 | 15 | BT000790 | BT000790 Arabidops |
| 43 | 508 | 25.0 | 1230 | 15 | AF428375 | AF428375 Arabidops |
| 44 | 508 | 25.0 | 1236 | 15 | AF083738 | AF083738 Arabidops |
| 45 | 508 | 25.0 | 1237 | 15 | AY087772 | AY087772 Arabidops |

ALIGNMENTS

| | | | | | | |
|------------|---|---|---------|------|--------|-----------------|
| RESULT 1 | BT000874 | Arabidopsis thaliana clone C104984 putative NAM (no apical meristem) protein (Atg02450) mRNA, complete cds. | 1171 bp | mRNA | linear | PLN 16-OCT-2002 |
| LOCUS | BT000874 | Arabidopsis thaliana (thale cress) | | | | |
| DEFINITION | BT000874.1 | GI:24030185 | | | | |
| ACCESSION | BT000874 | | | | | |
| VERSION | FLI CDNA. | | | | | |
| KEYWORDS | Arabidopsis thaliana | | | | | |
| SOURCE | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi | | | | | |
| ORGANISM | 1 (bases 1 to 1171) | | | | | |
| REFERENCE | Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R., and Theologis, A. | | | | | |
| AUTHORS | Arabidopsis Open Reading Frame (ORF) Clones | | | | | |
| JOURNAL | Unpublished | | | | | |
| REFERENCE | 2 (bases 1 to 1171) | | | | | |
| AUTHORS | Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R., and Theologis, A. | | | | | |

TITLE Direct Submission
JOURNAL Submitted (16-OCT-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
COMMENT Annotation based on January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES
Location/Qualifiers
1..1171
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="2"
/clone="C104984"
/ecotype="Columbia"
/note="this clone is in PUN1 51."

gene
1..1171
/gene="At2g02450"

CDS
1..1140
/gene="At2g02450"
/codon_start=1
/evidence="experimental"
/product="putative NAM (no apical meristem) protein"
/protein_id="AA041274.1"
/db_xref="GI:24030186"
/translation="MAIVSSTTSIIIPMSNOVNNKEGIEDNHRGGQESHVONEDAD
DHDHVMGPRFPHTEBELIEFYLRKVEGKRFNVELITFDLDLYDPWELPMAAI
GEKWFYVPRDKYRNGDRPNRTVSGYWKATGADMRISRTSPIGLKKTLVYSG
KAPKRTSWINNEIRLPHHETEKYKAEISLCRVYKRPVEDHPSVPSRSTRHNH
NSSTSRRLAQOQHSSSHSNNNNNNNNINLEKLSYEYSGDSTTTTTTNSNS
DVTIALQNIYRPMPTDTSNNTLVSTRNHQDDTETAVDDQLRVNYQISDGAATL
MPOTQALAMNMI PAGTIPNNALWDMNPVDPGNRDRHTNIPPK"
1141..1171
/gene="At2g02450"

3'UTR
misc_difference 1150
/gene="At2g02450"
/note="compared to genomic sequence"
/replace="a"

ORIGIN
Alignment Scores:
Pred. No.: 6,85e-150 Length: 1171
Score: 2034.00 Matches: 379
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 15 Gaps: 0

US-10-780-703-2 (1-379) x BT000874 (1-1171)

QY 1 MetAlaIleValSerThrThrSerIleIleProMetSerAsnGlnValAsnAsn 20
DB 1 ATGGCAATTGTATCTCTCCACACAGCATCATTCCTCCATGAGTAACCAAGTCAACATAAC 60
QY 21 GluLysGlyIleGluAspAsnAspHisArgGlyGlyGlnGluSerHisValGlnAsnGlu 40
DB 61 GAAAGAGGTATAGACACATGATCATAGAGGGCCCAAGAGAGTATCTCCAAATGAA 120
QY 41 AspGluAlaAspAspHisAspMetValMetProGlyPheArgPheHisProThr 60
DB 121 GATGAAGCTGATGATCATGATCATGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 61 GluGluGluLeuIleGluPheTyrLeuArgArgLysValGluGlyLysArgPheAsnVal 80
DB 181 GAAGAAGAACTCATAGAGTTTACCTTCCTCCGAAAAGTTTGAAGGCAACAGCGTTTAAATGTA 240
QY 81 GluLeuIleThrPheLeuAspLeuTyrArgTyrAspProTyrGluLeuProAlaMetAla 100
DB 241 GAACATCATCATCTTCTCTCATCTTATCTGCTATGATCTTCTGGAACTTCTGCTATGGCG 300
QY 101 AlaIleGlyGluLysGluTyrPheTyrValProArgAspArgLysTyrArgAsnGly 120
DB 301 GCGATAGGAGAGAAAGAGTGGTACTTCTATGTCGCAAGAGATCGGAAATATAGAAATGGA 360
QY 121 AspArgProAsnArgValThrThrSerGlyTyrTrpLysAlaThrGlyAlaAspArgMet 140

DB 361 GATAGACCGAACCGAGTAAACGACTTCAGATATTGGAAAGCCACCGAGGCTGATAGGATG 420
QY 141 IleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTyrSerGly 160
DB 421 ATCAGATCGAGAGACTTCTCGGCTATCGGATTAAGAAACCCCTAGTTTCTTACTCTGT 480
QY 161 LysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeuProHis 180
DB 481 AAAGCCCTTAAAGGCACCTCGTACTAGTTGGATCATGAAGAGTATCGTCTTCGACCAT 540
QY 181 GluThrGluLysTyrGlnLysAlaGluLeuSerLeuCysArgValTyrLysArgProGly 200
DB 541 GAAACCGAAGAGTACCAAAAGGCTGAAATATCATTTGTGCGAGTGTACAAAAGGCCAGCA 600
QY 201 ValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsnHisAsnSer 220
DB 601 GTAGAAGATCATCCATCGTACCAAGTCTCTCTCCCAAGACATCATAAACATTAACATCA 660
QY 221 SerThrSerArgLeuAlaLeuArgGlnGlnHisHisSerSerSerSerSerSerSer 240
DB 661 TCGACATCATCCGTTTACGCTTAAAGACAACAACCAACCATTCATCTCTCTCTTAATCAT 720
QY 241 SerAspAsnLeuAsnAsnAsnAsnAsnAsnAsnLeuGluLysLeuSerThrGlu 260
DB 721 TCCGACACACCTTAAACAACAACAACAACAACAATCTCGAGAAGCTCTCCACCGAA 780
QY 261 TyrSerGlyAspGlySerThrThrThrThrThrThrThrThrThrThrThrThrThr 280
DB 781 TATTCCGCGAGCGGAGCAGCAACAACAACAACAACAACAAGTAACTCTGACGTTTACCATT 840
QY 281 AlaLeuAlaAsnGlnAsnIleTyrArgProMetProTyrAspThrSerAsnAsnThrLeu 300
DB 841 GCTCTAGCCCAATCAAAACATATATCGTCAATGCTTACGACCAAGCAACCAACACATTG 900
QY 301 IleValSerThrArgAsnHisGlnAspAspGluThrAlaIleValAspAspLeuGln 320
DB 901 ATAGTCTCTACGAGAAATCATCAAGCGATGATGAACCTGCCATTGTTGACGATCTTCAA 960
QY 321 ArgLeuValAsnTyrGlnIleSerAspGlyAlaThrThrLeuMetProGlnThrGlnAla 340
DB 961 AGACTAGTTAACTACCAATATCATGATGAGCGAGCAACGCTAATGCTCTCAAACTCAAGCG 1020
QY 341 AlaLeuAlaMetAsnMetIleProAlaGlyThrIleProAsnAsnAlaLeuTrpAspMet 360
DB 1021 GCGTTAGCTATGAACATGATCTCTGAGAAAGATTCCAACATGCTTTTGGGAATG 1080
QY 361 TrpAsnProIleValProAspGlyAsnArgAspHisTyrThrAsnIleProPheLys 379
DB 1081 TGGATCCCAATAGTACCAAGTGAACAGAGATCACTATATATATCTCTTTTAAG 1137

RESULT 2
AC005312 74903 bp DNA linear PLN 11-MAR-2002
LOCUS Arabidopsis thaliana chromosome 2 clone T16F16 map C1C11A04,
DEFINITION complete sequence.
AC005312
VERSION AC005312.3 GI:20197334
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 74903)
AUTHORS Rounsley,S.D., Lin,X., Kaul,S., Shea,T.P., Fujii,C.Y., Mason,T.M.,
Shen,M., Renning,C.M., Fraser,C.M., Somerville,C.R. and Venter,J.C.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 74903)
AUTHORS Lin,X.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA


```

/product="putative NAM protein"
/db_xref="GI:21450867"
/translation="LYRDPWELPMAAIGKEKEWYFVPRDKYRNGRDRPRVTTSGY
WKATGADRMISERSRPGLKKTLLVYSGKAPKGRTRSWINNEVRLPHETEKYQKAE
ISLCVYRPGVEDSPSRSLSTRHHNHSSTSRSLALRQOHHSSSNHSNHNLLN
NNINLEKLSLEYSGDSTTTTNTNSNDVTIALANQNIYRPMPTDTSNNTLIVSTR
NHODDETAI VDDLRQVYQISDCAATLMPQTAALAMNMI PACTI PNNALWDMNP
IVPDGNRDHYTNIIPK"
882..1062
3'UTR
/gene="At2g02450"

ORIGIN
Alignment Scores:
Pred. No.: 6 12e-114 Length: 1062
Score: 1574.00 Matches: 292
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 77.4% Indels: 0
DB: 15 Gaps: 0

US-10-780-703-2 (1-379) x AF360308 (1-1062)

Qy 88 LeuTyrArgTyrAspProTyrGluLeuProAlaMetAlaAlaIleGlyGluLysGluTyr 107
Db 3 CTTTATCGTATGATCTCTGGCACTTCTGCTATGCGCGGATAGGAGAGAAAGAGTGG 62
Qy 108 TyrPheTyrValProArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThr 127
Db 63 TACTTCTATGTGCCAAGAGATCGGAATATAGAAATAGAGATAGACCGAACCGAGTAACG 122
Qy 128 ThrSerGlyTyrTrpLeuAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArg 147
Db 123 ACTTCAGGATATTGAAAGCCACCGAGCTGATAGATGATCGATCGAGACTTCTCGG 182
Qy 148 ProIleGlyLeuLysLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArg 167
Db 183 CCTATCGGATTAAGAAACCCCTAGTTTCTACTCTGGTAAAGCCCTTAAGGCACCTCGT 242
Qy 168 ThrSerTrpIleMetAsnGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLys 187
Db 243 ACTAGTTGGATCATGAACGAGTATGCTCTTCGCAACCATGAACCGAGAGTACCAAAAG 302
Qy 188 AlaGluIleSerLeuCyseArgValTyrLysArgProGlyValGluAspHisProSerVal 207
Db 303 GCTGAATATCATTTGCGCGAGTGACAAAGGCCAGGAGTAGAAGATCATCCATCGGTA 362
Qy 208 ProArgSerLeuSerThrArgHisHisAsnHisAsnSerSerThrSerArgLeuAla 227
Db 363 CCACGTTCTCTCCACAGACATCATACCATTAATCATCGACATCATCCGTTAGCC 422
Qy 228 LeuArgGlnGlnGlnHisSerSerSerSerSerHisSerAspAsnLeuAsnAsn 247
Db 423 TTAAGACAACAACAACACCATTCCTCTCTAATCATTCGTCGACAAACACCTTAACAAC 482
Qy 248 AsnAsnAsnIleAsnAsnLeuLysLeuSerThrGluTyrSerGlyAspGlySerThr 267
Db 483 AACACAACATCAACAATCTCGAGAGCTCTCCACCGAATATTCGCGGACGCGAGCACA 542
Qy 268 ThrThrThrThrThrAsnSerAsnSerAspValThrIleAlaLeuAlaAsnGlnAsnIle 287
Db 543 ACAACAACGACCAACAACAGTACTCTGAGCTTACCATTCGTCGACCAATCAACAACATA 602
Qy 288 TyrArgProMetProTyrAspThrSerAsnAsnThrLeuIleValSerThrArgAsnHis 307
Db 603 TATCGTCAATGCCCTTAGCACAACAAGCAACAACATTCATGATGATCTCTACGAGAATCAT 662
Qy 308 GlnAspAspGluThrAlaIleValAspAspLeuGlnArgLeuValAsnTyrGlnIle 327
Db 663 CAAGACGATGATGAACATGCGCATTTGTCAGCATCTTCAAAAGACTAGTTAACTTACCAATA 722
Qy 328 SerAspGlyAlaThrThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsnMetIle 347

```

```

Db 723 TCAGATGGAGCGACACACGCTAATGCTCAAACTCAAGCGCGTTAGCTATGACCATGATT 782
Qy 348 ProAlaGlyThrIleProAsnAsnAlaLeuTrpAspMetTrpAsnProIleValProAsp 367
Db 783 CCTCAGGAACGATTCACAAACATGCTTTGGGATATGTGGATATCAATAGTACCATGAT 842
Qy 368 GlyAsnArgAspHisTyrThrAsnIleProPheLys 379
Db 843 GGAACACAGAGATCACTATATATTCCTTTTAAG 878

AKI02902 1631 bp mRNA linear PLN 24-JUL-2003
Oryza sativa (japonica cultivar-group) cDNA clone:J033113D13, full
insert sequence.
AKI02902
AKI02902.1 GI:32988111
FLI_CDNA; CAP trapper.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1
The Rice Full-Length cDNA Consortium, National Institute of
Agricultural Sciences Rice Full-Length cDNA Project Team.;
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
Ohtsuki, K., Shihiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group; Otomo, Y., Murakami, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN;
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oeato, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M., and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)
12869764
2 (bases 1 to 1631)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kawagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Namiki, T., Narikawa, R., Niikura, J., Nishii, K., Nomura, K.,
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
Oeato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Shihiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yanagi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
Yoshimura, A.
Direct Submission
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
Agricultural Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kamondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)

```


Kodama, T., Kurotsaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Teunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayaishi, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ootani, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, P., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.
Location/Qualifiers
1. .2424
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J013151L09"

FEATURES

source

ORIGIN

Alignment Scores:
Pred. No.: 4,378-52 Length: 2424
Score: 789.50 Matches: 175
Percent Similarity: 55.7% Conservatives: 46
Best Local Similarity: 44.1% Mismatches: 75
Query Match: 38.8% Indels: 101
DB: 15 Gaps: 14

US-10-780-703-2 (1-379) x AK068393 (1-2424)

QY 44 AspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThrGluGluGlu 63
DB 135 GACGGGACGGGACGACATGTGTGTCGGGGTTCGGGTTCACCCGAGGAGGAG 194
QY 64 LeuileGluPheThrLeuArgArgLysValGluGlyLysArgPheAsnValGluLeuile 83
DB 195 CTCATCGAGTTCTACCTCCGGGAAAGTGGAGGGAAGCGGTTCAACATCGAGTCATC 254
QY 84 ThrPheLeuAspLeuThrArgThrArgProThrGluLeuProAlaMetAlaLeuileGly 103
DB 255 GCCTTCGTGCGACTCTACCGCTATGACCGCTGGGATCTTCCCGCTCGCCTCAATTGGG 314
QY 104 GluLysGluThrPheThrValProArgAspArgLysThrArgAsnGlyAspArgPro 123
DB 315 GACAGAGTGGTTCCTTCATGTGCGAGGAGCCGCAAGTACCAACCGCGATCGGCC 374
QY 124 AsnArgValThrThrSerGlyThrTrpLysAlaThrGlyAlaAspArgMetileArgSer 143
DB 375 ACCCGAGTACCGCATCGGGTACTCGAAGGCCACAGGGCGGATAGGTGTGAGGTG 434
QY 144 GluThrSerArgProLleGlyLysLysThrLeuValPheThrSerGlyLysAlaPro 163
DB 435 GAGGGTGACCGTCTATCGCTCAAGAGAGCGCTCTTCTACGTCCGCAAGGCGGCC 494
QY 164 LysGlyThrArgThrSerTrpIleMetAsnGluThrArgLeuProHisGluThrGlu 183
DB 495 AAGGGGCTTCGCGAGCGAGTGGATCATGAACAGAGTATGCTCTCCCTCATGCCGCGTAC 554
QY 184 LysThrGlnLysAlaGluSerLeuCysArgValThrLysArgProGlyValGluAsp 203
DB 555 CGCTACCAAG---GAAATTCGTGTGTCGAGTCTACAAAGCCCGCAGGATCGAGAC 611
QY 204 His-----ProSerValProArgSerLeu 211
DB 612 AACTTCCACCTCAGCGGCACAAACCACTAAATCTCCGGCTCAAAAGCGCGCGCGATG 671

QY 212 SerThrArgHis-----HisAsnHisAsnSerSerThrSerArgLeuAlaLeuArg 229
DB 672 GCGAAGAAGACACGCGCGCGAACCAGGAGTCAATCGAGCGCGCGCTCGCGCGCGCC 731
QY 230 GlnGlnGln-----HisHis-----Ser 235
DB 732 CTCGCCCGCATGTTTCGACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 791
QY 236 SerSerSerHisSerAspAsnAsnLeuAsnAsnAsnAsnAsnAsnAsnAsnAsn 251
DB 792 GCAACATACAGCATACACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 851
QY 251 ----- 251
DB 852 TCG 911
QY 252 ---AsnAsnLeuGluLysSerThrThrGluThrSerGlyAsp----- 264
DB 912 GCCTCGCTGAGCTCCACCACGCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 971
QY 265 ---GlySerThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 283
DB 972 GCG 1007
QY 284 AsnGlnAsnIleTyrArgProMetProTyrAspThr-----SerAsnAsn 298
DB 1008 CATCACCATCATCATCAGCAGCTGCTGCTTCACTCATGCTCTGCTCAACACCACTCG 1067
QY 299 ThrLeuLeuValSerThrArgAsnHisGlnAspAspGluThrAlaLeuValAspAsp 318
DB 1068 TCAGCAATGGCG-----ACCATCCCATCGACGAG 1097
QY 319 LeuGlnArgLeuVal-----AsnTyrGlnIleSer 328
DB 1098 CTGAGTAGGGGATTGGATCTTACAGCAACATCCACCATCAACCAACTCTACTACA 1157
QY 329 AspGlyAlaThrThrLeuMetProGlnThrGlnAlaLeuAlaMetAsnMetIlePro 348
DB 1158 CATCAATCACTGCTTTCGCGCATCGCAACAAAGCCCTCTT-----CTTCCA 1205
QY 349 AlaGlyThrIleProAsnAlaLeuTrpAspMetTrpAsnProIleVal 365
DB 1206 TTCGTAGCATG-----GAGAAGATTGGGAT---TGAACCTCTCTCTA 1247
RESULT 8
AY672069 1406 bp mRNA linear PLN 25-SEP-2004
LOCUS Hordeum vulgare subsp. vulgare NAC transcription factor mRNA,
DEFINITION complete cds.
ACCESSION AY672069
VERSION AY672069.1 GI:52352766
KEYWORDS
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
REFERENCE 1. (bases 1 to 1406)
AUTHORS Robertson, M.
TITLE Two Transcription Factors Are Negative Regulators of Gibberellin
Response in the HvSPY-Signaling Pathway in Barley Aleurone
Plant Physiol. 136 (1), 2747-2761 (2004)
JOURNAL
PUBMED 15347799
AUTHORS Robertson, M.
TITLE Direct Submission
JOURNAL Submitted (29-JUN-2004) Plant Industry, CSIRO, GPO Box 1600,
Canberra, ACT 2601, Australia
FEATURES Location/Qualifiers
1. .1406
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Himalaya"


```

/sub_species="vulgare"
/db_xref="taxon:112509"
4..1179
/notes="HSINAC; HvSPY-interacting NAC"
/codon_start=1
/product="NAC transcription factor"
/protein_id="AAU43824.1"
/db_xref="GI:52352767"
/translation="MTRSRPTTTMGSSVPDQHHQQOQHDGVDDGGLOHGGERHETVM
PGRFPHTEELIEFYLRKVDGKRFNDIADVDLYRDPDLPALASIGDKWEFFY
VPRDRYRNDNRNPTSPGKATGADRMVKVEGNSIGLKTIVFYVYKAPGLR
SSWIMNEYRPHQETERYKEISLCRVYKRPDDNLHLGTTTRSSGSRARTRST
AAHRTSAVATHROQPEVEGAGHSSALKAYNAHTTQGTNAASTWTSLSAAGVTP
AMFRTSAVASLSTSTSEEDCTSLVHHHLKGNHNPVMDLHSSHTAALLNTNSSAMA
TIPIELSLRAIGSYSQASNPENQPTAPMQGPLLNPSPLEKIDWNPILLESFKVCTSPK"

ORIGIN
Alignment Scores:
Pred. No.: 7,1e-52 Length: 1406
Score: 783.00 Matches: 181
Percent Similarity: 57.3% Conservative: 54
Best Local Similarity: 44.1% Mismatches: 82
Query Match: 38.5% Indels: 93
DB: 15 Gaps: 15

US-10-780-703-2 (1-379) x AY672069 (1-1406)

QY 21 GluLysGlyLeuGluAspAsnAspHisArgGlyGly-----Gln-GluSerHisVa 37
DB 3 GATGACTCGATCAAGGCGGACGACCAACCATGGGGGATGATGCGCAGACCAACACCA 62
QY 37 LGINAsnGluaspGluAlaasp-----AspHis-----AspHis-- 48
DB 63 GCAGCAGCAGCAGCAGCGGGAAGTGGACGCTGGGCGACCTGCAGCGCGGCGGACGCT 122
QY 49 -AspMetValMetProGlyPheArgPheHisProThrGluGluGluLeuLeuGluPheTy 68
DB 123 GGAGACGGTGATGCCCGGGTTCGTTTCCACCCGACGCGAGGAGGAGCTGATCGAGTTCTA 182
QY 68 rLeuArgArgLysValGluGlyLysArgPheAsnValGluLeuLeuThrPheLeuAspLe 88
DB 183 CCTTCGTCGCAAGTGGACGCGAAGCGCTTCAACATCGACTCATCGCTCCGTCGACCT 242
QY 88 uTyArgTyArgProTrpGluLeuProAlaMetAlaAlaLeuGlyGlyGluTrpTy 108
DB 243 CTACCGCTACGACCATGGGATCTCCCGCGACTAGCGTGCATCGGGACAGGAGTGTT 302
QY 108 rPheTyValProArgAspArgLysTyArgAsnGlyAspArgProAsnArgValThrTh 128
DB 303 CTTCTAGCTTCTCGGACCGCAAGTAGTACCGGACCGGACCGCGGACGAGGAGTGAGCC 362
QY 128 rSerGlyTyTrpLysAlaThrGlyAlaAspArgMetIleArg---SerGluThrSerAr 147
DB 363 GTCAGGGTACTGGAAGCAACGCGGGCGGACAGGATGTTGAAAGTCTGTGGAGGCAACCG 422
QY 147 gProIleGlyLeuLysLysThrLeuValPheTySerGlyLysAlaProLysGlyThrAr 167
DB 423 CTCCATCGGCTCAAGAAGACGCTCGCTTCTACGTTGGCAAGGACACCAAGGACTCCG 482
QY 167 gThrSerTrpIleMetAsnGluTyArgLeuProHisHisGluThrGluLysTyTrGlnLy 187
DB 483 CAGCAGCTGATCATGATGACGAGTACCCCTCGCCACACGCGTGAACACCGAGCTACCAAA 542
QY 187 sAlaGluIleSerLeuCyAsArgValTyLysArgProGlyValGluAspHis----- 204
DB 543 G---GAATTTTCGCTCTCGGGGCTATAAAGCGCCAGGATTCAGCACAACCTTCACCT 599
QY 205 -----ProSerValProArgSerLeuSerThrArgHisHisAsnHi 218
DB 600 CACCGGCACACACAAAGGTCGTCTGCTCCAGAGCGCGGACACGAGGACATACAGC 659
QY 218 sAsnSerSerThrSerArgLeuAlaLeuArgGlnGlnGln----- 232

```

```

660 AGCACACCGGAGCTCAGCGGTGGCAACTCACCGCCAGCAGCCGCCAGTCTTTGTCGAAGG 719
233 -----HisHisSerSerSerSerAsnHisSerAspAsnAsnLeuAsnAsnAsnAs 250
720 CCGCGGTTCACCACTCATCTCTGCTCTCAAGCGGTACAAC----- 759
250 nileAsnAsnLeuGluLysLeuSerThrGluTySerGlyAspGlySerThrThrThr 270
760 -----GCCACACGACCCA 773
270 rThrThrAsnSerAspValThrIleAlaLeuAlaAsnGlnAsnIleTyArgPr 290
774 AGGAACAATGCGACCGACCATGACA---TCACCTGTCAGCGGAGGAGTGCAGCCACC 830
290 oMetProTyArgThrSerAsnAsnThrLeuIleValSerThrArgAsnHisGlnAspAs 310
831 GCGGATGTTCCGGTCGACGGCTCTGTAGCCTCGCTGAGCTCTACGAATTCGACAGAGA 890
310 pAspGluThrAlaIleValAspAspLeuGlnArgLeuValAsnTyArgGlnIleSerAspG 330
891 GATGGCAGCTGCTCTACCAACACCGCTGAGGGGACACAC-----AACCCAGC 941
330 yAlaThrThrLeuMetProGlnThrGlnAlaAlaLeu-----AlaMe 344
942 AGTAATGCAACTGCATCTTCCACTCAGCTGCTGCTCAACACCAATTCCTCGGCAAT 1001
344 tAsnMetIleProAlaGlyThrIle----- 352
1002 GCGGACCATCTCCGATCGACGAGCTGAGTGGGGGATGGGTCTACAGTCAAGCTTCAAA 1061
353 -ProAsn-----AsnAl 356
1062 CCTAACACGCGCGGCGCAATGCGGCGCCATTCGTTAACTTCCTAGCTTGGAGAA 1121
356 aLeuTrpAspMetTrpAsnProIleVal 365
1122 GATCTGGGAC---TGGAACTCTCTCCTA 1146

RESULT 9
AX654050 1212 bp DNA linear PAT 22-MAR-2003
LOCUS Sequence 3920 from Patent WO03000898.
DEFINITION AX654050
ACCESSION AX654050.1 GI:29156864
VERSION
KEYWORDS
SOURCE
ORIGIN
REFERENCE
AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
Katsigiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
TITLE Plant genes involved in defense against pathogens
JOURNAL Patent: WO 03000898-A 3920 03-JAN-2003;
Syngenta Participations AG (CH)
FEATURES
source
1..1212
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"

ORIGIN
Alignment Scores:
Pred. No.: 8,45e-52 Length: 1212
Score: 781.00 Matches: 185
Percent Similarity: 57.1% Conservative: 47
Best Local Similarity: 45.6% Mismatches: 94
Query Match: 38.4% Indels: 81
DB: 6 Gaps: 11

US-10-780-703-2 (1-379) x AX654050 (1-1212)

```

QY 52 MetProGlyPheArgPheHisProThrGluGluLeuLeuLeuGluPheTyrLeuArgArg 71
Db 1 ATCCCGGGTTACAGTTCCAGCCGACGAGGAGAGTGTATCGAGTTCTACCTCCGCGG 60
QY 72 LysValGluGlyLysArgPheAsnValGluLeuLeuLeuPheLeuAspLeuTyrArgTyr 91
Db 61 AAGGTGGAGGGCAAGCGCTTCAACGTCGAGCTCATCACCCTTCCTCGACCTCTACCGCTAC 120
QY 92 AspProTyrGluLeuPro-----Ala 98
Db 121 GACCCCTGGAGCTCCCGGAATTGAGCATCTAATTAAACAAGCTCGATGGGGTGCAGCA 180
QY 99 MetAlaAlaGlyGlyLysGluTyrPheTyrValProArgAspArgLysTyrArg 118
Db 181 ATGCGCGCATAGGGGAGAGAGTGTCTTCTACGTCCTCGGACAGGAAGTACAGG 240
QY 119 AsnGlyAspArgProAsnArgValThrSerGlyTyrTriPylsAlaThrGlyAlaAsp 138
Db 241 AACGGGGACCGCGCAACCGGTGACGGGTCTGGGTACTGGAAGCGGCGGCGCGAC 300
QY 139 ArgMetIleArgSerGluThrSerArgProIleGlyLeuLysValThrLeuValPheTyr 158
Db 301 CGATGATCCGAGCGGAGAACACCGCCCATCGGGCTCAAGAAGACGCTTGTCTCTAC 360
QY 159 SerGlyLeuAlaProLysGlyThrArgThrSerTriPylsMetAsnGluTyrArgLeuPro 178
Db 361 TCGGCAAGCGCCCAAGGCGTCCGACGAGTGGATCATGAAGAGTATCGCTTCCC 420
QY 179 HisHisGluThrGluLysTyrGlnLysAlaGluLeuSerLeuCysArgValTyrLysArg 198
Db 421 CCGCGGACACCGCGCTACCAACAGACTGAAATCTCATATGCGCGGTGTACAGCGC 480
QY 199 ProGlyValGluAsp---HisProSerVal-ProArg-----Se 210
Db 481 ACCGATCGACGACGCGCATCGCCAGGTATTCACGCGCGGTGTCGCGCGCATCCCGC 540
QY 210 rLeuSerThrArgHisHisAsnHisAsnSerSerThrSerArgLeuAlaLeuArgG1 230
Db 541 GCGCGCGCGCGCGCGGTACAGGACACAGCAAGGTCTCTCGTCCGAGTCCAGCGCG 500
QY 230 n-----GlnGlnHisHisSerSerSerSerSerSerSerSerSerSerSerSer 247
Db 601 AGCGCGCTCCCACTCCGTCGCAAGCTCCACCTCTCAGCAGCGGAGTGCAGCGCGCGC 660
QY 247 AsnAsnAsnIleAsnAsnLeuGluLysLeuSerThrGluTyrSerGlyAspGlySerTh 267
Db 661 GCCATCGTCACGACACCGCGCCATGTTGGGCGCAACAGGACCGTCCGCGCGC----- 715
QY 267 rThrThrThrThrThrAsnSerSerSerSerSerSerSerSerSerSerSerSerSer 287
Db 716 -ACCATCAGCAGCAGCAGTCCAGCTGCGCAAGCTTGGCGGCTACCTGCGAGAT 774
QY 287 eTyrArg-----ProMetProTyrAspThrSerAsnAsnThrLeuLeuValSerThrAr 305
Db 775 TCCTCGATGCTTCGCTCCGCGCGCGATCAGCAGCAGCAATTTCCAGCAAGACTTCGCG 834
QY 305 g----- 305
Db 835 GGGCATTTGTACACGAGTACTCCAGAACACAGCGGTGCTTCGCTCCACGTAAGTACG 894
QY 306 -----AsnHisGlnAspAspAspGluThrAlaLeuValAspAspLe 319
Db 895 CTGCTCAACCTCGTCAACGCGCATCCATG-GGCAGCTCCGCGGTGCGCATTTGATGAGCT 953
QY 319 uGlnArgLeuValAsnTyr-----GlnIleSerAspGlyAlaThr----- 333
Db 954 GAGCTCGTGTGGGCGCACGCGCGCTTACATTAAACCGCGCGCGCGCGCACCA 1013
QY 334 -----LeuMetProGlnThrGlnAlaLeuAlaMetAsnMetIlePr 348
Db 1014 CTATAGTCAATTCCTTCACTGCTACGACGCGCGTCTGCGCATCAACCAACCGCGCGC 1073
QY 348 oAlaGly-----ThrIlePro-----As 354

Db 1074 GCTCGGACGACGCTC 1133
QY 354 nAsnAlaLeuTrpAspMetTrpAsnProIleValProAspGlyAsnArgAspHisTyrTh 374
Db 1134 CGACAGGATCTGGGACTGGAATAATCCGATCCCGAGGCGGAGGTAGGATTACAGCAC 1193
QY 374 rAsnIleProPheLys 379
Db 1194 CAGTACTGGATTCAG 1209
RESULT 10
AP007649 80577 bp DNA linear HTG 28-DEC-2004
LOCUS
DEFINITION
Lotus corniculatus var. japonicus chromosome 4 clone LjT35A11, ***
SEQUENCING IN PROGRESS ***, 15 unordered pieces.
ACCESSION
AP007649.1 GI:56805960
VERSION
HTG; HTGS PHASE1.
KEYWORDS
SOURCE
ORGANISM
Lotus corniculatus var. japonicus (Lotus japonicus)
Lotus corniculatus var. japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.
1
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Kaneko, T., Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
Structural Analysis of a Lotus japonicus Genome. XI. Sequence
Features and Mapping of Nine hundred twenty-one TAC Clones
Unpublished
2 (bases 1 to 80577)
Sato, S.
Direct Submission
Submitted (28-OCT-2004) Shusei Sato, Kazusa DNA Research Institute,
Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,
Chiba, 292-0818, Japan (E-mail: ssato@kazusa.or.jp,
URL: http://www.kazusa.or.jp/, Tel: 81-438-52-3935 (ex.2337),
Fax: 81-438-52-3934)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1062: contig of 1062 bp in length
* 1063 1162: gap of unknown length
* 1163 2101: contig of 939 bp in length
* 2102 2201: gap of unknown length
* 2202 3131: contig of 929 bp in length
* 3131 3230: gap of unknown length
* 3231 4436: contig of 1106 bp in length
* 4437 4336: gap of unknown length
* 4437 6509: contig of 1973 bp in length
* 6410 6510: gap of unknown length
* 6510 8323: contig of 1814 bp in length
* 8324 8423: gap of unknown length
* 8424 11366: contig of 2943 bp in length
* 11367 11466: gap of unknown length
* 11467 14883: contig of 3017 bp in length
* 14884 14583: gap of unknown length
* 14584 18560: contig of 3977 bp in length
* 18561 18660: gap of unknown length
* 18661 25852: contig of 7192 bp in length
* 25853 33903: contig of 7951 bp in length
* 33904 34003: gap of unknown length
* 34004 42262: contig of 8259 bp in length
* 42263 53643: contig of 11281 bp in length
* 53644 53743: gap of unknown length
* 53744 70427: contig of 16684 bp in length

QY 186 ----- 186
 Db 41929 TTTAAGTAGCAAAATTTAGATCCAAAGCTAGCTATAGCATGCCAACACAGTAGGTGGTG 41988
 QY 186 ----- 186
 Db 41989 AACTCGCAGCCACCATTAAATTTATTTCAAGAGTAGTCTTTTGGCATTTAATAAATACT 42048
 QY 186 ----- 186
 Db 42049 GATAGAAAAGATTTTCAGGAAATGTTTCAGCTTTGAGCACCGTTTGAATTAATTAATCTCA 42108
 QY 186 ----- 186
 Db 42109 CTCGAACATATTTATGATTACACGCAAGATTGTGAATTCGACATCATGTCTCGGAC 42168
 QY 186 ----- 186
 Db 42169 TTTGATCAATGATGATAGTCTAGACATTTCTCAATTTTCATCTCTAGTTACATATATTC 42228
 QY 187 ----- 187
 Db 42229 ATTGATTGATTGAATGTGAATGTGAATTTGGATCAATAGAGGCTGAGATATCGTCTTG 42289
 QY 193 sargValTyLysArgProGlyValGluAspHisProSerValProArgSerLeuSerTh 213
 Db 42289 CCGGGTATACAGAGAGCTGGAGTAGAGATCATCTTCCCTCCCTGGTGTGCTCCCAAT 42348
 QY 213 rArgHisAsnHisAsnSerSerThrSerSerArgLeuAlaLeuArgGlnGlnH 233
 Db 42349 TACCAAGCAGCCATCATCTCAAGATCATCAAAACCAAGTGAATAGCTCAGCATCAATA 42408
 QY 233 s-----HisSerSerSerSerSerSerSerSerSerSerSerSerSerSer 243
 Db 42409 TCTGCCACACATCATGGGCTTATTTCTTGGCCCAATCAATCAATCAATCAATCAATCAAT 42468
 QY 243 nAsnLeuAsnAsnAsnAsnAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 263
 Db 42469 ACAATAGTTGAGAT-----CATGAACGGAGGAGGATTTAGCAGACTGGTTCGGG 42519
 QY 263 yAspGlySerThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 283
 Db 42520 TTCCGGTTCAGAACATCATCTAACAACA-----GCTCTTGGGCTCTC 42561
 QY 283 aAsnGlnAsnIleTyArgProMetProTyArgProMetProTyArgProMetProTyArg 299
 Db 42562 CAAGTACAATGCTTACCGTGTGCGGTGGGTTCACCTCAGCAACACAAATACATAGGACT 42621
 QY 300 -----LeuIleValSerThrArgAsnHisGln----- 308
 Db 42622 GCCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 42681
 QY 308 ----- 308
 Db 42682 GCAACCAACCAATACTTTTAGTCCATGCTTTTCCCTGCAATGCTGCTGCTTCAAA 42741
 QY 309 -----AspAspAspGluThrAlaIleValAspAspLeuGlnArgLeuValAsnTyGln 326
 Db 42742 TAATAACAATAATAATAGTAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 42801
 QY 327 -----IleSerAspGlyAlaThrThrLeuMetProGlnTh 338
 Db 42802 ACAGTCTTCAATGTTCAAGTCAATCCCATCATCAATTAATTAATTAATTAATTAATTAAT 42861
 QY 338 rGlnAlaAlaLeuAlaMetAsnMetIleProAlaGlyThrIlePro-----AsnAs 355
 Db 42862 GCGCCACCAATGTCTCACTCACTACCA-----AATACATCTTCAACTACCTTCTCTGA 42918
 QY 355 nAlaLeuTrpAspMetTrpAspProIleValProAspGlyAsnArgAspHisTyThrAs 375
 Db 42919 CCGACTGTGGGAC-----TGAAATCCAATC-----CTGAGGTAACCAACGACCAACCAAT 42972
 QY 375 nIleProPheLys 379

Db 42973 CATGCCCTTCAAG 42985
 RESULT 12
 LOCUS AX654570
 DEFINITION Sequence 4440 from Patent WO03000898.
 ACCESSION AX654570
 VERSION AX654570.1 GI:29157384
 KEYWORDS
 SOURCE
 ORGANISM
 Oryza sativa
 Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE
 AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
 Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
 TITLE Plant genes involved in defense against pathogens
 JOURNAL Patent: WO '03000898-A 4440 03-JAN-2003;
 Syngenta Participations AG (CH)
 FEATURES
 source
 1..1250
 /organism="Oryza sativa"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4530"
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,638-50 Length: 1250
 Score: 759.00 Matches: 174
 Percent Similarity: 50.8% Conservative: 45
 Best Local Similarity: 40.4% Mismatches: 72
 Query Match: 37.3% Indels: 140
 DB: Gaps: 14
 US-10-780-703-2 (1-379) x AX654570 (1-1250)
 QY 50 MetValMetProGlyPheArgPheHisProThrGluGluGluLeuIleGluPheTyLeu 69
 Db 1 ATGGTAATGCGGGGTTCGGTTCACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60
 QY 70 ArgArgLysValGluGlyArgPheAsnValGluLeuIleThrPheLeuAspLeuTy 89
 Db 61 CGCGCAAGAGTGGAGGGGAAAGCGTTCAACATCGAGCTCATCGCTTCGTCGACCTCTAC 120
 QY 90 ArgTyArgProTrpGluLeuProAlaMetAlaIleGlyGlyGlyGlyTrpTyPhe 109
 Db 121 CGCTATGACCGGTGGGATCTTCCGCTCTGGCTCAATTTGGGACAAAGGAGTGGTCTTC 180
 QY 110 TyrValProArgAspArgLysTyArgAsnGlyAspArgProAsnArgValThrThrSer 129
 Db 181 TATGTGCAAGGAGCGCAAGTACCGAAACCGCGATCGGCCACCGAGTACCGCATCG 240
 QY 130 GlyTyTrpLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIle 149
 Db 241 GGGTACTGGAGGCCACAGGGCGGATAGATGGTGAAGTGGAGGAGGAGGAGGAGGAGGAGGAG 300
 QY 150 GlyLeuLysLysThrLeuValPheTySerGlyLysAlaProLysGlyThrArgThrSer 169
 Db 301 GGCCTCAAGAGACGCTCGTCTTCTACGTGGCAAGGGCGCCCAAGGGGCTTCGACGACG 360
 QY 170 TrpIleMetAsnGluTyArgLeuProHisHisGluThrGluLysTyGlnLys----- 187
 Db 361 TGGATCATGAACAGATATCGTCTCCCTCATGCGCGAGCTGACCGCTACCAAAAGATAGCT 420
 QY 187 ----- 187
 Db 421 TTTGATATAGTACCGCGGAGGAGCAAGCATATGTACTGTCATATGCTATTGCGAGCTCT 480
 QY 188 -----AlaGlu 189
 Db 481 TGTCTCTCAGCTGGTGATGAATTAAGTACACACCATCGGAAATTTGTATGTAGAGGAA 540

| | | | | | |
|--------------|---------|---------|--------------|----------|----------|
| AP008207_021 | 2100001 | 2210000 | AP008207_094 | 9400001 | 9510000 |
| AP008207_022 | 2200001 | 2310000 | AP008207_095 | 9500001 | 9610000 |
| AP008207_023 | 2300001 | 2400000 | AP008207_096 | 9600001 | 9710000 |
| AP008207_024 | 2400001 | 2510000 | AP008207_097 | 9700001 | 9810000 |
| AP008207_025 | 2500001 | 2610000 | AP008207_098 | 9800001 | 9910000 |
| AP008207_026 | 2600001 | 2710000 | AP008207_099 | 9900001 | 10010000 |
| AP008207_027 | 2700001 | 2810000 | AP008207_100 | 10000001 | 10110000 |
| AP008207_028 | 2800001 | 2910000 | AP008207_101 | 10100001 | 10210000 |
| AP008207_029 | 2900001 | 3010000 | AP008207_102 | 10200001 | 10310000 |
| AP008207_030 | 3000001 | 3110000 | AP008207_103 | 10300001 | 10410000 |
| AP008207_031 | 3100001 | 3210000 | AP008207_104 | 10400001 | 10510000 |
| AP008207_032 | 3200001 | 3310000 | AP008207_105 | 10500001 | 10610000 |
| AP008207_033 | 3300001 | 3410000 | AP008207_106 | 10600001 | 10710000 |
| AP008207_034 | 3400001 | 3510000 | AP008207_107 | 10700001 | 10810000 |
| AP008207_035 | 3500001 | 3610000 | AP008207_108 | 10800001 | 10910000 |
| AP008207_036 | 3600001 | 3710000 | AP008207_109 | 10900001 | 11010000 |
| AP008207_037 | 3700001 | 3810000 | AP008207_110 | 11000001 | 11110000 |
| AP008207_038 | 3800001 | 3910000 | AP008207_111 | 11100001 | 11210000 |
| AP008207_039 | 3900001 | 4010000 | AP008207_112 | 11200001 | 11310000 |
| AP008207_040 | 4000001 | 4110000 | AP008207_113 | 11300001 | 11410000 |
| AP008207_041 | 4100001 | 4210000 | AP008207_114 | 11400001 | 11510000 |
| AP008207_042 | 4200001 | 4310000 | AP008207_115 | 11500001 | 11610000 |
| AP008207_043 | 4300001 | 4410000 | AP008207_116 | 11600001 | 11710000 |
| AP008207_044 | 4400001 | 4510000 | AP008207_117 | 11700001 | 11810000 |
| AP008207_045 | 4500001 | 4610000 | AP008207_118 | 11800001 | 11910000 |
| AP008207_046 | 4600001 | 4710000 | AP008207_119 | 11900001 | 12010000 |
| AP008207_047 | 4700001 | 4810000 | AP008207_120 | 12000001 | 12110000 |
| AP008207_048 | 4800001 | 4910000 | AP008207_121 | 12100001 | 12210000 |
| AP008207_049 | 4900001 | 5010000 | AP008207_122 | 12200001 | 12310000 |
| AP008207_050 | 5000001 | 5110000 | AP008207_123 | 12300001 | 12410000 |
| AP008207_051 | 5100001 | 5210000 | AP008207_124 | 12400001 | 12510000 |
| AP008207_052 | 5200001 | 5310000 | AP008207_125 | 12500001 | 12610000 |
| AP008207_053 | 5300001 | 5410000 | AP008207_126 | 12600001 | 12710000 |
| AP008207_054 | 5400001 | 5510000 | AP008207_127 | 12700001 | 12810000 |
| AP008207_055 | 5500001 | 5610000 | AP008207_128 | 12800001 | 12910000 |
| AP008207_056 | 5600001 | 5710000 | AP008207_129 | 12900001 | 13010000 |
| AP008207_057 | 5700001 | 5810000 | AP008207_130 | 13000001 | 13110000 |
| AP008207_058 | 5800001 | 5910000 | AP008207_131 | 13100001 | 13210000 |
| AP008207_059 | 5900001 | 6010000 | AP008207_132 | 13200001 | 13310000 |
| AP008207_060 | 6000001 | 6110000 | AP008207_133 | 13300001 | 13410000 |
| AP008207_061 | 6100001 | 6210000 | AP008207_134 | 13400001 | 13510000 |
| AP008207_062 | 6200001 | 6310000 | AP008207_135 | 13500001 | 13610000 |
| AP008207_063 | 6300001 | 6410000 | AP008207_136 | 13600001 | 13710000 |
| AP008207_064 | 6400001 | 6510000 | AP008207_137 | 13700001 | 13810000 |
| AP008207_065 | 6500001 | 6610000 | AP008207_138 | 13800001 | 13910000 |
| AP008207_066 | 6600001 | 6710000 | AP008207_139 | 13900001 | 14010000 |
| AP008207_067 | 6700001 | 6810000 | AP008207_140 | 14000001 | 14110000 |
| AP008207_068 | 6800001 | 6910000 | AP008207_141 | 14100001 | 14210000 |
| AP008207_069 | 6900001 | 7010000 | AP008207_142 | 14200001 | 14310000 |
| AP008207_070 | 7000001 | 7110000 | AP008207_143 | 14300001 | 14410000 |
| AP008207_071 | 7100001 | 7210000 | AP008207_144 | 14400001 | 14510000 |
| AP008207_072 | 7200001 | 7310000 | AP008207_145 | 14500001 | 14610000 |
| AP008207_073 | 7300001 | 7410000 | AP008207_146 | 14600001 | 14710000 |
| AP008207_074 | 7400001 | 7510000 | AP008207_147 | 14700001 | 14810000 |
| AP008207_075 | 7500001 | 7610000 | AP008207_148 | 14800001 | 14910000 |
| AP008207_076 | 7600001 | 7710000 | AP008207_149 | 14900001 | 15010000 |
| AP008207_077 | 7700001 | 7810000 | AP008207_150 | 15000001 | 15110000 |
| AP008207_078 | 7800001 | 7910000 | AP008207_151 | 15100001 | 15210000 |
| AP008207_079 | 7900001 | 8010000 | AP008207_152 | 15200001 | 15310000 |
| AP008207_080 | 8000001 | 8110000 | AP008207_153 | 15300001 | 15410000 |
| AP008207_081 | 8100001 | 8210000 | AP008207_154 | 15400001 | 15510000 |
| AP008207_082 | 8200001 | 8310000 | AP008207_155 | 15500001 | 15610000 |
| AP008207_083 | 8300001 | 8410000 | AP008207_156 | 15600001 | 15710000 |
| AP008207_084 | 8400001 | 8510000 | AP008207_157 | 15700001 | 15810000 |
| AP008207_085 | 8500001 | 8610000 | AP008207_158 | 15800001 | 15910000 |
| AP008207_086 | 8600001 | 8710000 | AP008207_159 | 15900001 | 16010000 |
| AP008207_087 | 8700001 | 8810000 | AP008207_160 | 16000001 | 16110000 |
| AP008207_088 | 8800001 | 8910000 | AP008207_161 | 16100001 | 16210000 |
| AP008207_089 | 8900001 | 9010000 | AP008207_162 | 16200001 | 16310000 |
| AP008207_090 | 9000001 | 9110000 | AP008207_163 | 16300001 | 16410000 |
| AP008207_091 | 9100001 | 9210000 | AP008207_164 | 16400001 | 16510000 |
| AP008207_092 | 9200001 | 9310000 | AP008207_165 | 16500001 | 16610000 |
| AP008207_093 | 9300001 | 9410000 | AP008207_166 | 16600001 | 16710000 |

| | | | | | | |
|--------------|----------|----------|---|----------|---|--------|
| AP008207_126 | 12600001 | 12710000 | Pred. No.: | 7.67e-41 | Length: | 110000 |
| AP008207_127 | 12700001 | 12810000 | Score: | 672.50 | Matches: | 189 |
| AP008207_128 | 12800001 | 12910000 | Percent Similarity: | 39.9% | Conservative: | 49 |
| AP008207_129 | 12900001 | 13010000 | Best Local Similarity: | 31.7% | Mismatches: | 98 |
| AP008207_130 | 13000001 | 13110000 | Query Match: | 33.1% | Indels: | 263 |
| AP008207_131 | 13100001 | 13210000 | DB: | 15 | Gaps: | 12 |
| AP008207_132 | 13200001 | 13310000 | US-10-780-703-2 (1-379) x AP008207_386 (1-110000) | | | |
| AP008207_133 | 13300001 | 13410000 | QY | 42 | GlualaAspHisAspHisMetValMetProGlyPheArgPheHisProThrGlu | 61 |
| AP008207_134 | 13400001 | 13510000 | DB | 5196 | GAGGTAGATGGGCATGAGGATGACCTGGGTGATGCCCGGTTCAGGTTCACCCGACGGAG | 5255 |
| AP008207_135 | 13500001 | 13610000 | QY | 62 | GlulGluLeuIleGluPheTyrLeuArgArgLysValGluGlyLysArgPheAsnValGlu | 81 |
| AP008207_136 | 13600001 | 13710000 | DB | 5256 | GAGGAGCTGATCGAGTTCTACCTCCGCGGAGGTGGAGGCAAGCGCTTCAACGTCGAG | 5315 |
| AP008207_137 | 13700001 | 13810000 | QY | 82 | LeuileThrPheLeuAspLeuTyrArgTyrAspProTyrGluLeuPro | 97 |
| AP008207_138 | 13800001 | 13910000 | DB | 5316 | CTCATCACCTTCTCGACCTTACCGCTACGACCCCTGGGAGCT-CGCCGGTAATTAGTT | 5374 |
| AP008207_139 | 13900001 | 14010000 | QY | 97 | ----- | 97 |
| AP008207_140 | 14000001 | 14110000 | DB | 5375 | AGTTACTCTCCGCTCGATCATATCAATCGATTGATTTAGGAATTGAGCATCTAATTAAC | 5434 |
| AP008207_141 | 14100001 | 14210000 | QY | 98 | -----AlaMetAlaAlaIleGlyGluLysGluTyrPheTyrVal | 111 |
| AP008207_142 | 14200001 | 14310000 | DB | 5435 | AAGCTCGATGGGGTGCAGCAATGGCGGATAGGGGAGAGAGAGTGTTCCTACGTG | 5494 |
| AP008207_143 | 14300001 | 14410000 | QY | 112 | ProArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyr | 131 |
| AP008207_144 | 14400001 | 14510000 | DB | 5495 | CCTCGGGACAGGAAGTACAGGAACGGGACCGGCCGACCGGTGACGGGTGCGGGTAC | 5554 |
| AP008207_145 | 14500001 | 14610000 | QY | 132 | TyrLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeu | 151 |
| AP008207_146 | 14600001 | 14710000 | DB | 5555 | TGGAAGGCGACGGGGCGCCGCGATGATCCGAGCCGAGAACACCGCCCATTCGGGCTC | 5614 |
| AP008207_147 | 14700001 | 14810000 | QY | 152 | LysLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTyrPile | 171 |
| AP008207_148 | 14800001 | 14910000 | DB | 5615 | ARGAAGACGCTTGTCTTCTTACTCCGGCAGGGCCCGCCAGGGCGTCCGACAGCTGATC | 5674 |
| AP008207_149 | 14900001 | 15010000 | QY | 172 | MetAsnGluTyrArgLeuProHisGluThrGluLysTyrGlnLys | 187 |
| AP008207_150 | 15000001 | 15110000 | DB | 5675 | ATGAACGAGTATCGCTTCCCGCCGCGACACCGCGCTACCAACAGGTACCCATCCAT | 5734 |
| AP008207_151 | 15100001 | 15210000 | QY | 187 | ----- | 187 |
| AP008207_152 | 15200001 | 15310000 | DB | 5735 | CCATGAATCTACTCGCGCTTCGTTCTCTCGATCGATCTCTCTCTCATCAATTAATT | 5794 |
| AP008207_153 | 15300001 | 15410000 | QY | 187 | ----- | 187 |
| AP008207_154 | 15400001 | 15510000 | DB | 5795 | TGTAGGGGATTGAATTCTCAGTACTAGGTAGAGGATCTTGTGTGTGTGATTCTGCT | 5854 |
| AP008207_155 | 15500001 | 15610000 | QY | 187 | ----- | 187 |
| AP008207_156 | 15600001 | 15710000 | DB | 5855 | TGCTTTTAGTTGCACTGCTAGCTGGCGCTCATGGTTGCCCATGAATTTAGCGGTAGTA | 5914 |
| AP008207_157 | 15700001 | 15810000 | QY | 187 | ----- | 187 |
| AP008207_158 | 15800001 | 15910000 | DB | 5915 | TTTAAGAGTGGGTATGATTAGGGTTTCATGAATTTCTTTCTTTGGGTGCTGATTTTTT | 5974 |
| AP008207_159 | 15900001 | 16010000 | QY | 187 | ----- | 187 |
| AP008207_160 | 16000001 | 16110000 | DB | 5975 | TGGGGTGTTCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT | 6034 |
| AP008207_161 | 16100001 | 16210000 | QY | 187 | ----- | 187 |
| AP008207_162 | 16200001 | 16310000 | DB | 6035 | TCGATCCCTAGATGCTATATATCTTGAATATATATATATATATATATATATATATATAT | 6094 |
| AP008207_163 | 16300001 | 16410000 | QY | 187 | ----- | 187 |
| AP008207_164 | 16400001 | 16510000 | DB | 6095 | TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAATTCAGTAATTCAGTA | 6154 |
| AP008207_165 | 16500001 | 16610000 | QY | 187 | ----- | 187 |
| AP008207_166 | 16600001 | 16710000 | DB | 6095 | TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAATTCAGTAATTCAGTA | 6154 |
| AP008207_167 | 16700001 | 16810000 | QY | 187 | ----- | 187 |
| AP008207_168 | 16800001 | 16910000 | DB | 6095 | TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAATTCAGTAATTCAGTA | 6154 |
| AP008207_169 | 16900001 | 17010000 | QY | 187 | ----- | 187 |
| AP008207_170 | 17000001 | 17110000 | DB | 6095 | TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAATTCAGTAATTCAGTA | 6154 |
| AP008207_171 | 17100001 | 17210000 | QY | 187 | ----- | 187 |
| AP008207_172 | 17200001 | 17310000 | DB | 6095 | TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAATTCAGTAATTCAGTA | 6154 |
| AP008207_173 | 17300001 | 17410000 | QY | 187 | ----- | 187 |
| AP008207_174 | 17400001 | 17510000 | DB | 6095 | TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAATTCAGTAATTCAGTA | 6154 |
| AP008207_175 | 17500001 | 17610000 | QY | 187 | ----- | 187 |
| AP008207_176 | 17600001 | 17710000 | DB | 6095 | TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAATTCAGTAATTCAGTA | 6154 |
| AP008207_177 | 17700001 | 17810000 | QY | 187 | ----- | 187 |
| AP008207_178 | 17800001 | 17910000 | DB | 6095 | TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAATTCAGTAATTCAGTA | 6154 |
| AP008207_179 | 17900001 | 18010000 | QY | 187 | ----- | 187 |
| AP008207_180 | 18000001 | 18110000 | DB | 6095 | TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAATTCAGTAATTCAGTA | 6154 |
| AP008207_181 | 18100001 | 18210000 | QY | 187 | ----- | 187 |
| AP008207_182 | 18200001 | 18310000 | DB | 6095 | TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAATTCAGTAATTCAGTA | 6154 |
| AP008207_183 | 18300001 | 18410000 | QY | 187 | ----- | 187 |
| AP008207_184 | 18400001 | 18510000 | DB | 6095 | TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAATTCAGTAATTCAGTA | 6154 |
| AP008207_185 | 18500001 | 18610000 | QY | 187 | ----- | 187 |
| AP008207_186 | 18600001 | 18710000 | DB | 6095 | TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAATTCAGTAATTCAGTA | 6154 |
| AP008207_187 | 18700001 | 18810000 | QY | 187 | ----- | 187 |
| AP008207_188 | 18800001 | 18910000 | DB | 6095 | TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAATTCAGTAATTCAGTA | 6154 |
| AP008207_189 | 18900001 | 19010000 | QY | 187 | ----- | 187 |
| AP008207_190 | 19000001 | 19110000 | DB | 6095 | TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAATTCAGTAATTCAGTA | 6154 |
| AP008207_191 | 19100001 | 19210000 | QY | 187 | ----- | 187 |
| AP008207_192 | 19200001 | 19310000 | DB | 6095 | TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAATTCAGTAATTCAGTA | 6154 |
| AP008207_193 | 19300001 | 19410000 | QY | 187 | ----- | 187 |
| AP008207_194 | 19400001 | 19510000 | DB | 6095 | TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAATTCAGTAATTCAGTA | 6154 |
| AP008207_195 | 19500001 | 19610000 | QY | 187 | ----- | 187 |
| AP008207_196 | 19600001 | 19710000 | DB | 6095 | TCATTCTGAGCTGAATTTCTCACAACCTTTTCAGTAATTCAGTAATTCAGTAATTCAGTA | 6154 |

```
QY 188 -----Al 188
Db 6155 TGAATTTTCAGTGAATAATTAACCCCTATCTCTCGCGTCCATGAATTCAGAC 6214
QY 188 aGluLeuSerLeuCysArgValTyrIlyAspProGlyValGluAsp---HisProSerVa 207
Db 6215 TGAATATCTCACTATCGCGGTGTACAAAGCGCACCGCATCGACGCGCGCATGGCCAGGT 6274
QY 207 1-ProArg-----SerLeuSerThrArgHisHisAsnHisAsnS 220
Db 6275 ATCCACGGCGCGGTGCTCGGCGCACTCCCGCGCGCGCGCGCACCGGTACAGACAA 6334
QY 220 erSerThrSerSerArgLeuAlaLeuArgGln-----GlnGlnHisHisSerSers 237
Db 6335 CAAGCAAGGCTCTCTCGACGTCCAGCGCGCGCGCGCTCCCACTCCGTCCAAGTCCA 6394
QY 237 erSerAsnHisSerAspAsnAsnLeuAsnAsnAsnAsnAsnLeuAsnLeuLysL 257
Db 6395 CCTCCTCAGACGAGTGACGTCGCGCGCGCGCGCGCATCGTCACGAGCACCGCGCATG 6454
QY 257 euSerThrGluTyrSerGlyAspGlySerThrThrThrThrThrThrThrThrThr 277
Db 6455 GCGGCACAGGACCGCGTACCTGCAGATTCCTCGATGGCTTCGGTCCGCGCGCGCA 6508
QY 277 spValThrIleAlaLeuAlaAsnGlnAsnIleTyrArg-----ProMetProTyrAsp 295
Db 6509 TGCCAAAGCTTCGCGCGGTACTCTGCAGATTCCTCGATGGCTTCGGTCCGCGCGCG 6568
QY 295 hrSerAsnAsnThrLeuIleValSerThrArg----- 305
Db 6569 TCAGCAGCAGCAATTTTCAGCAAGACTTCGCGCGCGGATTGTACAGCAGTACTCCAAG 6628
QY 306 -----AsnHisGlnA 309
Db 6629 CACGAGCGGTGCTTCGCGCTCCACGTACTCGTGTCTCACTCGTCAACGCGCGCATCC 6688
QY 309 spAspAspGluThrAlaIleValAspAspLeuGlnArgLeuValAsnTyr----- 325
Db 6689 G-GGCAGCTCCGCGCTGCCATTGATGAGTGTGCTGGTGGCCACGCGCGCGT 6747
QY 326 --GlnIleSerAspGlyAlaThr-----LeuMetProGlnT 338
Db 6748 CCTACATTAAACCCCGCGCGCGCGCACCACTATAGTCAATTCTTCACCTGCTACGA 6807
QY 338 hrGlnAlaAlaLeuAlaMetAsnMetIleProAlaGly----- 350
Db 6808 CGCGCTCTCGCATCAACCAACGCGCGCGCGCTCGGACAGCAGCGCGCGCGCGG 6867
QY 351 --ThrIlePro-----AsnAsnAlaLeuTyrAspMetTyrAsnProI 364
Db 6868 CGAGCTGCGGATGTGCTCGCGCGCTTCTCGCAGCAGGATCTGGGACTGGATATATCCA 6927
QY 364 leValProAspGlyAsnArgAspHisTyrThrAsnIleProPheLys 379
Db 6928 TCCCGAGCGCGGAGGTAGGATTACAGCACCAGTACTGGATTCAAG 6974
```

Search completed: April 6, 2006, 11:43:26
Job time : 6538 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2006, 09:13:25 ; Search time 12942 Seconds
(without alignments)
11445.992 Million cell updates/sec

Title: US-10-780-703-3

Perfect score: 2606

Sequence: 1 atggcaattgtatctctcc.....ctaattcttttaagtaa 2606

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb ba.*
- 2: gb in.*
- 3: gb env.*
- 4: gb om.*
- 5: gb ov.*
- 6: gb pat.*
- 7: gb ph.*
- 8: gb pr.*
- 9: gb ro.*
- 10: gb sts.*
- 11: gb sy.*
- 12: gb un.*
- 13: gb vi.*
- 14: gb trg.*
- 15: gb pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 2606 | 100.0 | 74903 | 15 | AC005312 Arabidops |
| 2 | 476.2 | 18.3 | 534 | 10 | CR381764 Arabidops |
| 3 | 474.2 | 18.2 | 541 | 10 | CR381364 Arabidops |
| 4 | 467.2 | 17.9 | 1062 | 15 | AF360308 Arabidops |
| 5 | 467.2 | 17.9 | 1171 | 15 | BT000874 Arabidops |
| 6 | 463 | 17.8 | 524 | 10 | CR381358 Arabidops |
| 7 | 456.2 | 17.5 | 505 | 10 | CR381827 Arabidops |
| 8 | 447.8 | 17.2 | 530 | 10 | CR381634 Arabidops |
| 9 | 234 | 9.0 | 80577 | 14 | AP007649 Arabidops |
| 10 | 234 | 9.0 | 91718 | 14 | AP007409 Arabidops |
| 11 | 167.6 | 6.4 | 110000 | 15 | AP008207_385 |
| 12 | 167.6 | 6.4 | 110000 | 15 | AP008207_386 |
| 13 | 167.6 | 6.4 | 150064 | 15 | AP003346 Arabidops |
| 14 | 167.6 | 6.4 | 191022 | 15 | AP003431 Arabidops |
| 15 | 164 | 6.3 | 1212 | 6 | AX654050 Sequence |
| 16 | 158.6 | 6.1 | 1631 | 15 | AK102902 Oryza sat |
| 17 | 151.8 | 5.8 | 110000 | 15 | AP008214 007 |
| 18 | 151.8 | 5.8 | 185545 | 15 | AP005657 Oryza sat |

| | | | | | |
|------|-------|-----|--------|----|---------------------|
| 19 | 149.2 | 5.7 | 1640 | 15 | AK072275 Oryza sat |
| 20 | 149.2 | 5.7 | 1891 | 15 | AK071052 Oryza sat |
| 21 | 149.2 | 5.7 | 2424 | 15 | AK068393 Oryza sat |
| 22 | 149 | 5.7 | 1250 | 6 | AX654570 Sequence |
| c 23 | 135.8 | 5.2 | 110000 | 15 | AP008211_203 |
| c 24 | 135.8 | 5.2 | 173074 | 15 | AC137623 Oryza sat |
| 25 | 132.6 | 5.1 | 549 | 6 | AX653764 Sequence |
| 26 | 126.2 | 4.8 | 1406 | 15 | AY672069 Hordeum v |
| 27 | 120.8 | 4.6 | 202169 | 14 | AC149549 Gorilla g |
| 28 | 118.6 | 4.6 | 337 | 10 | AL808204 Arabidops |
| 29 | 118 | 4.5 | 1689 | 15 | AY974352 Glycine m |
| 30 | 117.2 | 4.5 | 1688 | 15 | AY974351 Glycine m |
| 31 | 117.2 | 4.5 | 82893 | 15 | AB022218 Arabidops |
| 32 | 117.2 | 4.5 | 82893 | 15 | AC024081 Arabidops |
| c 33 | 112.8 | 4.3 | 116862 | 14 | AC160836 Medicago |
| 34 | 112.2 | 4.3 | 945 | 6 | AX506848 Sequence |
| 35 | 112.2 | 4.3 | 945 | 6 | AX651806 Sequence |
| c 36 | 111.8 | 4.3 | 82381 | 15 | ATF27G19 Arabidops |
| c 37 | 111.8 | 4.3 | 200001 | 15 | ATCHRIV67 Arabidops |
| 38 | 111.6 | 4.3 | 1106 | 15 | DQ028771 Glycine m |
| 39 | 110.4 | 4.2 | 1095 | 15 | AB049071 Arabidops |
| 40 | 110.4 | 4.2 | 1126 | 15 | BT005044 Arabidops |
| 41 | 110.4 | 4.2 | 1319 | 6 | CS137982 Sequence |
| 42 | 110.4 | 4.2 | 1719 | 15 | BT004079 Arabidops |
| 43 | 110.2 | 4.2 | 1152 | 15 | DQ028772 Glycine m |
| c 44 | 109.2 | 4.2 | 44499 | 15 | AB016876 Arabidops |
| c 45 | 109 | 4.2 | 104204 | 15 | ATT9C5 Arabidops |

ALIGNMENTS

RESULT 1
AC005312
LOCUS Arabidopsis thaliana chromosome 2 clone T16F16 map C1C11A04, linear 74903 bp .DNA
DEFINITION complete sequence.

AC005312
AC005312.3 GI:20197334

VERSION HTG.

KEYWORDS Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1 (bases 1 to 74903)

AUTHORS Rounsley,S.D., Lin,X., Kaul,S., Shea,T.P., Fujii,C.Y., Mason,T.M., Shen,M., Rouning,C.M., Fraser,C.M., Somerville,C.R. and Venter,J.C.

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 74903)

AUTHORS Lin,X.

TITLE Direct Submission

JOURNAL Submitted (09-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA

REFERENCE 3 (bases 1 to 74903)

AUTHORS Town,C.D. and Kaul,S.

TITLE Direct Submission

JOURNAL Submitted (27-FEB-2002) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
On Apr 18, 2002 this sequence version replaced gi:6598458.

COMMENT Location/Qualifiers

FEATURES
source
1. 74903
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="2"
/map="C1C11A04"
/clone="T16F16"
/ecotype="Columbia"
complement (<1. 2731)
/genes="At2g02220"
/notes="synonym: T16F16.1"
complement (<1. 2731)

```

/genes="At2g02220"
complement (1..2731)
/genes="At2g02220"
/codon_start=1
/product="putative protein kinase"
/db_xref="GI:20197335"
/translation="MRVHRCVIVIELTELLCFYSESQTSRCHPHDLEALRDFIA
HLSPKPDGWINSSSDCCNWTGITNSNNTGVRILEGNKKLSGLSGLKDEI
RVNLGRNFIDKSIDIPISFNLKNLQTLIDSSNDLSGIPITSINLPALQFDSSNKN
GSPSHI CHNSTQIRVVKLAVNFAGNFTSGFKCVLLHLCIGMNDLGTNPEDLFH
LKNLNLGIGENLSLSIRNLSSLDVSNWLFSGEIPDVDFELPQLKFFLQ
TNGFIGIPKSLANSPLNLLNNSLSGLMLNCTAMALNSLDLGTNRNFRGLPE
NLPDCKLNVLNARNTFHGOVPEKPNFESLYFSLSNSSLANISSALGILQHCNKL
TLVLVTFHGEALPDSSSLHPEKLVVLVAVNCRLTGSMRSLSSNELQLDLNWR
LTGATFSGIDKALFDLUSNNSFTGEIPKSLTKLESLSRNSISNEFSPDPFPMK
RNESALQYNOIFGPTTIELGHNLSGPIWEEFNGNLKGLHVFIDKWNALSGIPSS
LSGWTLEALDNNRLSGIPVSLQSLFSKFSVYNNLSVIGPQGFQTFPNS
FESNLCEGRHFPCECTESALIKRRRRSGGDI GMAIGAFSGVFLTLTLLSVLRA
RRSGVDPEIESESNNRKEIGEIKLVLFQSDNKLSDLDLSTNSPDQANI
CGCGFVYKATLPDGKVAIKKLSGDCGQIEREPAEVETLSRAQHNVLVLRGCF
YKNDRLUTISYMEGNSLDYWHLRNDGPALLKWKRLRIAQGAAGLLYLHEGCDPHI
LHRDITSNLLDENFNSHLADFGLARLMSPYETHVSTDLVGTGLYIPPEYGA"
/genes="At2g02220"
/notes="overlap with BAC clone F504 (AC005936:1..1865)."
3226..3247
/rpt_family="AT_rich"
complement (3504..3524)
/rpt_family="AT_rich"
complement (4101..4133)
/rpt_family="AT_rich"
complement (6000..7202)
/genes="At2g02230"
/notes="synonym: T16F16.2"
complement (join(<6000..6237,6295..6555,6649..6767,
6867..7202))
/genes="At2g02230"
complement (join(6000..6237,6295..6555,6649..6767,
6867..7202))
/codon_start=1
/product="putative phloem-specific lectin"
/protein_id="AAC78508.1"
/db_xref="GI:3894158"
/translation="MSQIHGDSNGSGGGGSRNDEISVTRASRFDALPEDCISKVI
SHTSPRDACVVASVSKVSAQSDLVWEMFLPESEYSSLVQSAHLSKKEIPLSLAD
NSVLVGGKSFVWEKASGKKCYMLSAMELTIWGDSPAYKMTITVPESKFEKVALR
NVCWFVRGKISGMLSKGTHYSVVYVFTANGRSYGFDLVPVEAGVGFVKVATKS
VYFESGNADRSATSHYSGISEREVEGERGMNVGPKERVDSGSEVELGKFYIN
NGCGDGSDEIISIMETQNGNWKSLIIQGIIRPERSN"
/genes="At2g02240"
complement (7966..9209)
/notes="synonym: T16F16.3; pseudogene, phloem-specific
lectin"
/pseudo
complement (10262..11417)
/genes="At2g02250"
/notes="synonym: T16F16.4"
complement (join(<10262..10691,10775..10896,11052..>11417))
/genes="At2g02250"
complement (join(10262..10691,10775..10896,11052..11417))
/codon_start=1
/product="lectin-like protein"
/protein_id="AAC78509.1"
/db_xref="GI:3894159"
/translation="MGHKQSDVRGKRGKVPKVGSSMVQKRVESSGAILISGSLFDN
LPDSCISNIISFTSPRDACVAASVKTFESAVNSDVWDFLPSDYSSLVPPSRVFS
KRELFAICDNVLVDGGKSWLEKENGKCFMLSPKSKMTITVPTQYRWISIP
EARFEVPELNLVWCFVGRGMTKELSPGTRYSAIVYFKTKNCTNGLDVPVEATVG
LVQSESSORHI YFVGPSDQRDRDTRDVTPTKRRKDGWMEALGQFFNESGCDVDTS
ILBKIPTYKRGILLIQQIERPTKSLFYI"

```

```

complement (12061..13367)
/genes="At2g02260"
/notes="synonym: T16F16.5; pseudogene, reverse
transcriptase"
/pseudo
complement (13604..15132)
/genes="At2g02270"
/notes="synonym: T16F16.6"
complement (join(<13604..13663,13728..14100,14227..14312,
14854..>15132))
/genes="At2g02270"
complement (join(13604..13663,13728..14100,14227..14312,
14854..15132))
/codon_start=1
/product="putative phloem-specific lectin"
/protein_id="AAC78510.1"
/db_xref="GI:3894160"
/translation="MAQGEIVPRSPFDGLPENCISNIISFTTPRDACFAASVSKAFE
SAVQSDSVMEKFLPDYSSVFLSKKELCFSLCRVPLLIIEGKKSFWDKTSG
BKCIWLSKGMVVISNVSPOPEEVPEDLYDSHFEVCGRLSTKYLSPRTSYVIVFT
EVLVPGVILEPFPFRVPGFTDLKVEREYVTRPKREKKNWEALCEFFNETSCDV
EVSVDNSYKWSGLVIOGIEFRPTKKHQVKKNCIVNILRFYCNV"
/genes="At2g02280"
complement (16098..17046)
/notes="synonym: T16F16.7"
complement (join(<16098..16334,16849..>17046))
/genes="At2g02280"
complement (join(16098..16334,16849..17046))
/codon_start=1
/product="putative phloem-specific lectin"
/protein_id="AAC78511.1"
/db_xref="GI:3894161"
/translation="WNTQLSKTRYSAIVVTKTYRPHGFKHIGVGFHGTGTPKAKR
WERKLDNDLWCKCKFKSKQKQFYNNKSTFTDKPITHLKLEGGDMATEFGEF
FAEGGLDCCDEIVLSVIDIDYAYWKGLIIQGIIDIRPTKSP"
16391..16419
/rpt_family="AT_rich"
16523..16558
/rpt_family="AT_rich"
complement (16628..16649)
/rpt_family="(TA)n"
16672..16711
/rpt_family="(A)n"
complement (16770..16793)
/rpt_family="AT_rich"
complement (17280..17307)
/rpt_family="AT_rich"
complement (17616..17643)
/rpt_family="AT_rich"
complement (17834..17896)
/rpt_family="AT_rich"
complement (18077..18151)
/rpt_family="(GAA)n"
complement (18618..18659)
/rpt_family="AT_rich"
19173..19192
/rpt_family="(A)n"
20060..20968
/genes="At2g02290"
/notes="synonym: T16F16.8; predicted by genescan and
genefinder"
<20060..>20968
/genes="At2g02290"
20060..20968
/genes="At2g02290"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC78512.1"
/db_xref="GI:3894162"
/translation="MSLVENLSLEPQPKRKIDPEYKSPNSCSHWFRNKKVCASN
TIVDNYQGRSFDYLVGTIGMSNEALGFTKRLISQTSWLEDKHLVLDLDTLVHTIK

```


VSQSESEKYYITEEVESKDLRRFNTGFPPEESLIKLRSFVHQFLKECNEMFSLVYVTK
GGYDAQVLEIMIDPDKIYGNRVITRRSEPGFKTLDLVLADERGIVVVDDKSSWPH
DKNLLQIARYKYFQDQSCLLSECKKXIDESDEKPLNTALRFLMDVHEERPCDSRK
DLDSDVDRPLAKEISLRWKRKADGVV"
complement (22168..23432)
/gene="At2902300"
/note="synonym: T16F16.9"
complement (join(<22168..22558,22680..22798,23088..>23432))
/gene="At2902300"
complement (join(22168..22558,22680..22798,23088..23432))
/gene="At2902300"
/codon_start=1
/product="putative phloem-specific lectin"
/protein_id="AAC78513.1"
/db_xref="GI:3894163"

Query Match 100.0%; Score 2606; DB 15; Length 74903;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-------|--|-------|
| QY | 1 | ATGGCAATTGTATCTCCACAAAGCATCATTCCTCATGAGTAACCAAGTCAACAATAAC | 60 |
| DB | 63651 | ATGGCAATTGTATCTCCACAAAGCATCATTCCTCATGAGTAACCAAGTCAACAATAAC | 63710 |
| QY | 61 | GAAGAGGTATAGAGACAAATGATCATAGAGCGGCCAAGAGATCATGTCCAAATGAA | 120 |
| DB | 63711 | GAAGAGGTATAGAGACAAATGATCATAGAGCGGCCAAGAGATCATGTCCAAATGAA | 63770 |
| QY | 121 | GATGAAGCTGATGATCATGATCATGATGATGATGATGATGATGATGATGATGATGAT | 180 |
| DB | 63771 | GATGAAGCTGATGATCATGATCATGATGATGATGATGATGATGATGATGATGATGAT | 63830 |
| QY | 181 | GAAGAGAACTCATAGAGTTTACCTTCGCCGAAAGTTGAAGCAACGCTTTAAATGTA | 240 |
| DB | 63831 | GAAGAGAACTCATAGAGTTTACCTTCGCCGAAAGTTGAAGCAACGCTTTAAATGTA | 63890 |
| QY | 241 | GAATCATCACTTTCTCTGATCTTTATCGTATGATCTCTGGAACTTCTCGGTAAATAT | 300 |
| DB | 63891 | GAATCATCACTTTCTCTGATCTTTATCGTATGATCTCTGGAACTTCTCGGTAAATAT | 63950 |
| QY | 301 | ACATTACATAAACAACATAAATCATCTCAAACTATTGGAAATCTTAAATTTCTATCA | 360 |
| DB | 63951 | ACATTACATAAACAACATAAATCATCTCAAACTATTGGAAATCTTAAATTTCTATCA | 64010 |
| QY | 361 | TATGTTAAGATCTTCTCTCTCTATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT | 420 |
| DB | 64011 | TATGTTAAGATCTTCTCTCTCTCTATCATCTCTCTCTCTCTCTCTCTCTCTCTCT | 64070 |
| QY | 421 | ATATATGTACCTACCTCTTATGAAGTATTACTATGTGATCGTGTAAACATTTCTCAATAT | 480 |
| DB | 64071 | ATATATGTACCTACCTCTTATGAAGTATTACTATGTGATCGTGTAAACATTTCTCAATAT | 64130 |
| QY | 481 | CTTTAAAGCTTCT | 540 |
| DB | 64131 | CTTTAAAGCTTCT | 64190 |
| QY | 541 | TATATATCATAGATATACAAATATGTTATGTTTCTTAAATTAACCTTAAATTAACCTA | 600 |
| DB | 64191 | TATATATCATAGATATACAAATATGTTATGTTTCTTAAATTAACCTTAAATTAACCTA | 64250 |
| QY | 601 | TCATAGATATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT | 660 |
| DB | 64251 | TCATAGATATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT | 64310 |
| QY | 661 | GTGCCAAGAGATCGGAAATATAGAAATGGAGATAGACCGGAAACCGAGTAAACCTCAGGA | 720 |
| DB | 64311 | GTGCCAAGAGATCGGAAATATAGAAATGGAGATAGACCGGAAACCGAGTAAACCTCAGGA | 64370 |
| QY | 721 | TATTTGGAAGCCACCGAGCTGATAGATGATGATGATGATGATGATGATGATGATGATGAT | 780 |
| DB | 64371 | TATTTGGAAGCCACCGAGCTGATAGATGATGATGATGATGATGATGATGATGATGATGAT | 64430 |
| QY | 781 | TTAAAGAAAACCCCTAGTTTCTCTACTCTGGTAAAGCCCCCTAAAGGCATCTCGTACTAGTTGG | 840 |

| | | | |
|----|-------|--|-------|
| DB | 64431 | TTAAAGAAAACCCCTAGTTTCTCTACTCTGGTAAAGCCCCCTAAAGGCATCTCGTACTAGTTGG | 64490 |
| QY | 841 | ATCATGAACGAGTATCGCTTCCGCAACCATGAAACCGGAGAGTACCAAAAGGTATAAATTT | 900 |
| DB | 64491 | ATCATGAACGAGTATCGCTTCCGCAACCATGAAACCGGAGAGTACCAAAAGGTATAAATTT | 64550 |
| QY | 901 | CTACTATAAATCTATATATATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTAT | 960 |
| DB | 64551 | CTACTATAAATCTATATATATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTAT | 64610 |
| QY | 961 | TTAGAGCCCTTTAAATTTGAAATTAATCCCTAGACAGTTTGAATTTTCTTTTCTGACTAG | 1020 |
| DB | 64611 | TTAGAGCCCTTTAAATTTGAAATTAATCCCTAGACAGTTTGAATTTTCTTTTCTGACTAG | 64670 |
| QY | 1021 | TTTTTATTTATTTTGGAAATTTGATTCGATAGATCAAAAATATCTTGTGAATGGACTAA | 1080 |
| DB | 64671 | TTTTTATTTATTTTGGAAATTTGATTCGATAGATCAAAAATATCTTGTGAATGGACTAA | 64730 |
| QY | 1081 | ATGTCAGCGCGCTTTGGCTTTAAATCCAGAAAATGTTTCATGTCATATCGGTGAACCTCT | 1140 |
| DB | 64731 | ATGTCAGCGCGCTTTGGCTTTAAATCCAGAAAATGTTTCATGTCATATCGGTGAACCTCT | 64790 |
| QY | 1141 | TTAAATTTGCTAGACATGGCCCATATGTTTATAGTAGAATAACATTAATAGATAGATGCATAC | 1200 |
| DB | 64791 | TTAAATTTGCTAGACATGGCCCATATGTTTATAGTAGAATAACATTAATAGATAGATGCATAC | 64850 |
| QY | 1201 | ACATATATATAAACAACAGATATCACTCGACATTCATATACCTTAAATTTCTGCAGAGA | 1260 |
| DB | 64851 | ACATATATATAAACAACAGATATCACTCGACATTCATATACCTTAAATTTCTGCAGAGA | 64910 |
| QY | 1261 | CATAGTTAGTTTCTTCTTACAAATTTATGACATGAAATGTTCTCTCTCTCTCTCTCTCTCT | 1320 |
| DB | 64911 | CATAGTTAGTTTCTTCTTACAAATTTATGACATGAAATGTTCTCTCTCTCTCTCTCTCTCT | 64970 |
| QY | 1321 | CATGCTCTTCTTATTAAGTTTACCCCAACATTTTGTGAAATTAATTTGGCATATATGAATATA | 1380 |
| DB | 64971 | CATGCTCTTCTTATTAAGTTTACCCCAACATTTTGTGAAATTAATTTGGCATATATGAATATA | 65030 |
| QY | 1381 | CCAAACATATTTATATGCGAACATTTTAAATCTATACGAATGATACGGTTTATGGAGTAG | 1440 |
| DB | 65031 | CCAAACATATTTATATGCGAACATTTTAAATCTATACGAATGATACGGTTTATGGAGTAG | 65090 |
| QY | 1441 | ACCGAAAATAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT | 1500 |
| DB | 65091 | ACCGAAAATAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT | 65150 |
| QY | 1501 | TCCTTCGACTTATATGTCGTCACCATTTTGAACCATATAAATTTTATAAATTTTCTATGTAT | 1560 |
| DB | 65151 | TCCTTCGACTTATATGTCGTCACCATTTTGAACCATATAAATTTTATAAATTTTCTATGTAT | 65210 |
| QY | 1561 | ATATATGATATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT | 1620 |
| DB | 65211 | ATATATGATATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT | 65270 |
| QY | 1621 | TATGTCATGATTTGTCATCTTCTGTCATCTTATTAATTTTATTAATTTTATTAATTTTATTA | 1680 |
| DB | 65271 | TATGTCATGATTTGTCATCTTCTGTCATCTTATTAATTTTATTAATTTTATTAATTTTATTA | 65330 |
| QY | 1681 | ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 1740 |
| DB | 65331 | ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 65390 |
| QY | 1741 | ACTCTAAATTTCTTCTTAAACGTCATCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT | 1800 |
| DB | 65391 | ACTCTAAATTTCTTCTTAAACGTCATCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT | 65450 |
| QY | 1801 | TGAATCTTCTTCTTCAAAACCATATCTTCTGATAAATGATATTTTCTCATAGATATCTTCTGTC | 1860 |
| DB | 65451 | TGAATCTTCTTCTTCAAAACCATATCTTCTGATAAATGATATTTTCTCATAGATATCTTCTGTC | 65510 |
| QY | 1861 | TATATTTGATAAATTTTGATATATGATCAAGTCTCTAATCAATGTCTCATGTATTAATTTAT | 1920 |

| | | | | | |
|--|---|--|---------|------------|--|
| Db | 65511 | TATATTTGATAAATTTGATATATGATCAAGTCTCTAATCAATGCTCATGTATAATTAT | 65570 | COMMENT | AT denotes an activation tag dissociation transposon within a single line. ET an enhancer trap dissociation transposon, GT a gene trap dissociation transposon, MT a mis-expression enhancer trap dissociation transposon, SM a defective suppressor mutator transposon. _3 denotes a sequence derived from the 3' end of the transposon. _5 denotes a sequence derived from the 5' end of the transposon. BBRC GARNET, ATIS project |
| Qy | 1921 | AGCTGAATATATCTTGGCGAGTGTAACAAAGCCAGAGTAGAAGATCATCCATCGG | 1980 | | On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock code: N165702 |
| Db | 65571 | AGCTGAATATATCTTGGCGAGTGTAACAAAGCCAGAGTAGAAGATCATCCATCGG | 65630 | | |
| Qy | 1981 | TACACAGTCTCTCTCCACAAAGACATCATAAACATACTCATCGACATCATCCGTTTAC | 2040 | | |
| Db | 65631 | TACACAGTCTCTCTCCACAAAGACATCATAAACATACTCATCGACATCATCCGTTTAC | 65690 | | |
| Qy | 2041 | CCTTAAGACAAACAAACACATTCATCTCTCTTAATCATTCGACAAACCTTAACA | 2100 | | |
| Db | 65691 | CCTTAAGACAAACAAACACATTCATCTCTCTTAATCATTCGACAAACCTTAACA | 65750 | | |
| Qy | 2101 | ACAACACACATCAACATCTCGAGAGCTCTCCACCGAATATTCGGCGAGCGGACGA | 2160 | | |
| Db | 65751 | ACAACACACATCAACATCTCGAGAGCTCTCCACCGAATATTCGGCGAGCGGACGA | 65810 | | |
| Qy | 2161 | CAACAAACAGCACAAACAGATTAACCTGACGTTTACCATTTGCTTAGCCAAATCAAAACA | 2220 | | |
| Db | 65811 | CAACAAACAGCACAAACAGATTAACCTGACGTTTACCATTTGCTTAGCCAAATCAAAACA | 65870 | | |
| Qy | 2221 | TATATCGTCAATGCTTACGACAAAGCAACAAACATGATGCTTACGAGAAATC | 2280 | | |
| Db | 65871 | TATATCGTCAATGCTTACGACAAAGCAACAAACATGATGCTTACGAGAAATC | 65930 | | |
| Qy | 2281 | ATCAAGACGATGAATGCTTACGACAAAGCAACAAACATGATGCTTACGAGAAATC | 2340 | | |
| Db | 65931 | ATCAAGACGATGAATGCTTACGACAAAGCAACAAACATGATGCTTACGAGAAATC | 65990 | | |
| Qy | 2341 | TATCAGATGAGGTAAACATCAATCAACCAATCTTCAAAATGCTCAACAGTTTCATCAT | 2400 | | |
| Db | 65991 | TATCAGATGAGGTAAACATCAATCAACCAATCTTCAAAATGCTCAACAGTTTCATCAT | 66050 | | |
| Qy | 2401 | CTCAACAAACAAATGCTTACGACAAAGCAACAAATGCTTCAACAGTTTCATCAT | 2460 | | |
| Db | 66051 | CTCAACAAACAAATGCTTACGACAAAGCAACAAATGCTTCAACAGTTTCATCAT | 66110 | | |
| Qy | 2461 | CAACGTAATGCTTCAACCTCAAGCGCGTTAGTATGATGATGCTTCAACAGTTTCATCAT | 2520 | | |
| Db | 66111 | CAACGTAATGCTTCAACCTCAAGCGCGTTAGTATGATGATGCTTCAACAGTTTCATCAT | 66170 | | |
| Qy | 2521 | TTCCAAACAAATGCTTGGGATATGGAATCAATAGTACAGATGGAACAGAGATC | 2580 | | |
| Db | 66171 | TTCCAAACAAATGCTTGGGATATGGAATCAATAGTACAGATGGAACAGAGATC | 66230 | | |
| Qy | 2581 | ACTATACTAATATCTCTTTTAAAGTAA 2606 | | | |
| Db | 66231 | ACTATACTAATATCTCTTTTAAAGTAA 66256 | | | |
| RESULT 2 | | | | | |
| LOCUS | CR381764 | 534 bp | DNA | linear | STS 27-MAR-2004 |
| DEFINITION | Arabidopsis thaliana transposon insertion STS GT_5.102100, sequence tagged site. | | | | |
| ACCESSION | CR381764 | | | | |
| VERSION | CR381764.1 | GI:45773245 | | | |
| KEYWORDS | STS; STS, sequence tagged site. | | | | |
| SOURCE | Arabidopsis thaliana (thale cress) | | | | |
| ORGANISM | Arabidopsis thaliana | | | | |
| REFERENCE | | | | | |
| AUTHORS | Clarke, J.H., Bowles, B., Carter, J., Hart, D., McCullagh, B., Walsh, S., Latham, S., Legrys, C., Jones, J.D.G. and Bevan, M. | | | | |
| JOURNAL | Unpublished | | | | |
| REFERENCE | 2 (bases 1 to 534) | | | | |
| AUTHORS | Clarke, J.H. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (26-MAR-2004) Clarke J.H., John Innes Centre, Colney Lane, Norwich, NR4 7UJ, UK | | | | |
| FEATURES | | | | | |
| source | 1..534 | | | | |
| Location/Qualifiers | | | | | |
| /organism="Arabidopsis thaliana" | | | | | |
| /mol_type="genomic DNA" | | | | | |
| /db_xref="taxon:3702" | | | | | |
| /clone="AC005312" | | | | | |
| /ecotype="Landsberg erecta NASC stock code NW20" | | | | | |
| /note="Unknown" | | | | | |
| 1..534 | | | | | |
| /standard_name="GT_5.102100" | | | | | |
| ORIGIN | | | | | |
| Query Match | 18.3% | Score | 476.2 | DB 10 | Length 534 |
| Best Local Similarity | 97.7% | Pred. No. | 1.6e-68 | | |
| Matches | 515 | Conservative | 0 | Mismatches | 8 |
| | | | | Indels | 4 |
| | | | | Gaps | 3 |
| Qy | 956 | GGTGGTGAGCGCTTTAAATTAATGAAATTAATCCTAGACAGTTTGAATTTTTTTC-TTTTTT | 1014 | | |
| Db | 10 | GGTGGTGAGCGCTTTAAATTAATGAAATTAATCCTAGACAGTTTGAATTTTTTTC-TTTTTT | 69 | | |
| Qy | 1015 | GACTAGTTTTATTTATTTATTTTGGAAATGATTCGATAAGATCAAAATACCTTGTGAATG | 1074 | | |
| Db | 70 | GACTAGTTTTATTTATTTATTTTGGAAATGATTCGATAAGATCAAAATACCTTGTGAATG | 129 | | |
| Qy | 1075 | GACTAAATGTCAGCGCGCTTTGGCTTAAATCCAG-AAATATGTTTCATGTCATATCGCT | 1133 | | |
| Db | 130 | GACTAAATGTCAGCGCGCTTTGGCTTAAATCCAGAAATATGTTTCATGTCATATCGCT | 189 | | |
| Qy | 1134 | GAACTCTTTAAATTTGTCAGACATGGCCCATATGTTATAGTAGAATACATTAATAGATAGA | 1193 | | |
| Db | 190 | GAACTCTTTAAATTTGTCAGACATGGCCCATATGTTATAGTAGAATACATTAATAGATAGA | 249 | | |
| Qy | 1194 | TGCATACATATATATAAACAACACAAAGTATCACACTCGACATTCATATACCTTAATTC | 1253 | | |
| Db | 250 | TGCATACAC-ATATATAAACAACACAAAGTATCACACTCGACATTCATATACCTTAATTC | 307 | | |
| Qy | 1254 | GCAGAGACATAGTTAGTTTCTTCAATTTATGACATGAATGCTTCTCTCTCTCTCAC | 1313 | | |
| Db | 308 | GCAGAGACATAGTTAGTTTCTTCAATTTATGACATGAATGCTTCTCTCTCTCTCAC | 367 | | |
| Qy | 1314 | ATTAAATTCATGCTTCTTCTTAAAGTTACCCCAACATTTTGAATTAATTTGGCATATATG | 1373 | | |
| Db | 368 | ATTAAATTCATGCTTCTTCTTAAAGTTACCCCAACATTTTGAATTAATTTGGCATATATG | 427 | | |
| Qy | 1374 | AATTATACCAACATATTTATATCGCAACATTTAAATCTATACGAATGATAACGGTTTAT | 1433 | | |
| Db | 428 | AATTATACCAACATATCTATATCGCAACATTTAAATCTATACGAATGATAACGGTTTAT | 487 | | |
| Qy | 1434 | GGAGTAGACCGGAAAAAATATATGATATACGAAAAATGCAATGGATA 1480 | | | |
| Db | 488 | GGAGTAGACCGGAAAAAATATATGATATACGAAAAATGCAATGGATA 534 | | | |
| RESULT 3 | | | | | |
| LOCUS | CR381364 | 541 bp | DNA | linear | STS 27-MAR-2004 |
| DEFINITION | Arabidopsis thaliana transposon insertion STS GT_5.101351, sequence tagged site. | | | | |
| ACCESSION | CR381364 | | | | |
| VERSION | CR381364.1 | GI:45772845 | | | |
| KEYWORDS | STS; STS, sequence tagged site. | | | | |
| SOURCE | Arabidopsis thaliana (thale cress) | | | | |
| ORGANISM | Arabidopsis thaliana | | | | |
| REFERENCE | | | | | |
| AUTHORS | Clarke, J.H., Bowles, B., Carter, J., Hart, D., McCullagh, B., Walsh, S., Latham, S., Legrys, C., Jones, J.D.G. and Bevan, M. | | | | |
| JOURNAL | Unpublished | | | | |
| REFERENCE | 2 (bases 1 to 534) | | | | |
| AUTHORS | Clarke, J.H. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (26-MAR-2004) Clarke J.H., John Innes Centre, Colney Lane, Norwich, NR4 7UJ, UK | | | | |

spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 Clarke, J.H., Bowles, B., Carter, J., Hart, D., McCullagh, B., Walsh, S., Langham, S., LeGrys, C., Jones, J.D.G. and Bevan, M.

Unpublished

2 (bases 1 to 541)

Direct Submission

Submitted (26-MAR-2004)

Lane, Norwich, NR4 7UJ, UK

AT denotes an activation tag dissociation transposon within a

single line, Et an enhancer trap dissociation transposon, GT a gene

trap dissociation transposon, MT a mis-expression enhancer trap

dissociation transposon, SM a defective suppressor mutator

transposon. 3 denotes a sequence derived from the 3' end of the

transposon. 5 denotes a sequence derived from the 5' end of the

transposon. BSRSC GARNET, ARIS project

On-line seed stock requests: <http://nasc.nott.ac.uk/> NASC stock

code: N165059.

Location/Qualifiers

1. 541

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/db_xref="taxon:3702"

/clones="AC005312"

/ecotype="Landsberg erecta NASC stock code NW20"

/note="Unknown"

1. 541

/standard_name="GT_5.101351"

ST5

ORIGIN

Query Match 18.2%; Score 474.2; DB 10; Length 541;

Best Local Similarity 97.7%; Pred. No. 3.5e-68;

Matches 513; Conservative 0; Mismatches 8; Indels 4; Gaps 3;

956 GGTGGTGGGCGCTTTAAATTTGAATTAATCCCTAGACAGTTCGAATTTTTC-TTTTTT 1014

19 GGTGGTGGGCGCTTTAAATTTGAATTAATCCCTAGACAGTTCGAATTTTTC-TTTTTT 78

1015 GACTAGTTTATTTATTTATTTTGAATTTGATTCGATAAGATCAAAATCTTGGAATG 1074

79 GACTAGTTTATTTATTTATTTTGAATTTGATTCGATAAGATCAAAATCTTGGAATG 138

1075 GACTAATGTCAGCGCGTTTGGCTTAAATTCAG-AAAAATCTTCATGTCATATCGGT 1133

139 GACTAATGTCAGCGCGTTTGGCTTAAATTCAG-AAAAATCTTCATGTCATATCGGT 198

1134 GAACCTTTTAAATTTGCTAGACATGCGCCCATATGTTATAGTAGAATACATTAATAGATGA 1193

199 GAACCTTTTAAATTTGCTAGACATGCGCCCATATGTTATAGTAGAATACATTAATAGATGA 258

1194 TGCATACATATATATAAACAACAAGATATCACATTCGACATTCATATACCTTAAATCT 1253

259 TGCATACAC--ATATATAACAACAAGATATCACATTCGACATTCATATACCTTAAATCT 316

1254 GCAGAGCATAGTTAGTTTCTTACATTTATGACATGAATGTTCTGCTCTTCCTCAC 1313

317 GCAGAGCATAGTTAGTTTCTTACATTTATGACATGAATGTTCTGCTCTTCCTCAC 376

1314 ATTAATTCATGCTCTTCTATTAAGTTACCCCAACATTTTTTGAATAATTTGGCATATATG 1373

377 ATTAATTCATGCTCTTCTATTAAGTTACCCCAACATTTTTTGAATAATTTGGCATATAT 436

1374 AATTATACCAACATTTATATATGGAACATTTAAATCTATACGAATGATAACGGTTTAT 1433

437 AATTATACCAACATTTATATATGGAACATTTAAATCTATACGAATGATAACGGTTTAT 496

1434 GGAGTAGACCGAAAAATATTTATGATACGGAATGACAAATGACAAATGGA 1478

497 GGAGTAGACCGAAAAATATTTATGATACGGAATGACAAATGACAAATGGA 541

RESULT 4

LOCUS AF360308

DEFINITION Arabidopsis thaliana putative NAM protein (At2g02450) mRNA, partial

cds.

ACCESSION AF360308

VERSION AF360308.1

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana

Arabidopsis thaliana (thale cress)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 1062)

AUTHORS Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K., Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Arabisopsis Full Length cDNA Clones

Unpublished

2 (bases 1 to 1062)

Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K., Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Direct Submission

Submitted (13-MAR-2001)

Plant Gene Expression Center, 800 Buchanan

Street, Albany, CA 94710, USA

RIKEN Genomic Sciences Center (GSC) members carried out the

collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN

Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,

Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,

Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the

sequencing and annotation of the RAPL cDNAs: Yamada, K., Liu, S.X.,

Sakano, H., Pham, P.K., Banh, J., Chung, M.K., Goldsmith, A.D.,

Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L., Chen, H.,

Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J.,

Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A.,

Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to

this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC)

contributed equally to this work as PIs.

Annotatation is based on the January 2002 version of the Arabidopsis

genome submitted to Genbank.

Location/Qualifiers

1. 1062

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosomes="2"

/clones="RAFL09-17-121 (R09375)"

/ecotypes="Columbia"

/notes="This clone is in a modified pBluescript vector

(FLC-1) as a BamHI/XhoI insert."

<1. 1062

/gene="At2g02450"

<1. 881

/gene="At2g02450"

/note="putative no apical meristem protein"

/codon_start=3

/evidence="experimental"

/product="putative NAM protein"

/protein_id="AAK26018.2"

gene

CDS

AUTHORS Clarke, J.H.
 TITLE Direct Submission
 JOURNAL Submitted (26-MAR-2004) Clarke J.H., John Innes Centre, Colney Lane, Norwich, NR4 7UJ, UK

COMMENT AT denotes an activation tag dissociation transposon within a single line, ET an enhancer trap dissociation transposon, GT a gene trap dissociation transposon, MT a mis-expression enhancer trap dissociation transposon, SM a defective suppressor mutator transposon. _3 denotes a sequence derived from the 3' end of the transposon, _5 denotes a sequence derived from the 5' end of the transposon. BBSRC GARNET, ATIS project
 On-line seed stock requests: <http://nasc.nott.ac.uk/> NASC stock code: N165877.

FEATURES
 Location/Qualifiers
 . source
 1..505
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /db_xref="taxon:3702"
 /clone="AC005312"
 /ecotype="Landsberg erecta NASC stock code NW20"
 /note="Unknown"
 1..505
 /standard_name="GT_5.102321"

STS
 1..505
 /standard_name="GT_5.102321"

ORIGIN
 Query Match 17.5%; Score 456.2; DB 10; Length 505;
 Best Local Similarity 97.6%; Pred. No. 3.3e-65;
 Matches 495; Conservative 0; Mismatches 8; Indels 4; Gaps 3;

QY 964 GGCCTTTAAATGAAATTAATCCCTAGACAGTTGAAATTTTTC-TTTTTGACTAGTT 1022
 DB 1 GGCCTTTAAATGAAATTAATCCCTAGACAGTTGAAATTTTTC-TTTTTGACTAGTT 60

QY 1023 TTATTTATTTATTTGGAATGATTCGATAAGATCAAAAATATCTGTGAATGCACTAAAT 1082
 DB 61 TTATTTATTTATTTGGAATGATTCGATAAGATCAAAAATATCTGTGAATGCACTAAAT 120

QY 1083 GTCAGCGGGGTTTGGCTTAAATCCAG-AAAATGTTTCATGTCATATGGTGAACTCTT 1141
 DB 121 GTCAGCGGGGTTTGGCTTAAATCCAGAAAATGTTTCATGTCATATGGTGAACTCTT 180

QY 1142 TAAATGCTAGACATGCCCATATGTTATAGTAGATACATTAATAGATAGATGATACATA 1201
 DB 181 TAAATGCTAGACATGCCCATATGTTATAGTAGATACATTAATAGATAGATGATACATA 240

QY 1202 CATATATATAACACACAAGATGATCACCTGCACATTCATATACCTTAAATTCGACAGAC 1261
 DB 241 C--ATATATAACACACAAGATGATCACCTGCACATTCATATACCTTAAATTCGACAGAC 298

QY 1262 ATAGTTAGTTTCTTACAAATTTATGACATGAATGTTCCGCTCTTCCCTCACATTAATTC 1321
 DB 299 ATAGTTAGTTTCTTACAAATTTATGACATGAATGTTCCGCTCTTCCCTCACATTAATTC 358

QY 1322 ATGCTCTCTATTTAAAGTTACCAACATTTTGGAAATAATTTGGCATATATGAATTATAC 1381
 DB 359 ATGCTCTCTATTTAAAGTTACCAACATTTTGGAAATAATTTGGCATATATGAATTATAC 418

QY 1382 CAACATATTTATATGCGAACAATTTAAATCTATACGAATGATACCGTTTTATGGAGTAGA 1441
 DB 419 CAACATATCTATATGCGAACAATTTAAATCTATACGAAGATACCGTTTTATGGAGTAGA 478

QY 1442 CCGAAAAAATATTTATGTTATACGGAAAA 1468
 DB 479 CCGAAAAAATATCTATGTTATACGGATAA 505

RESULT 8
 CR381634
 LOCUS Arabidopsis thaliana transposon insertion STS GT_5.101721, sequence tagged site.
 DEFINITION CR381634
 ACCESSION CR381634
 VERSION CR381634.1 GI:45773115

KEYWORDS STS, sequence tagged site.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
 1 Clarke, J.H., Bowles, B., Carter, J., Hart, D., McCullagh, B., Walsh, S., Langham, S., LeGry, C., Jones, J.D.G. and Bevan, M.
 Unpublished
 2 (bases 1 to 530)
 Clarke, J.H.
 Direct Submission
 JOURNAL Submitted (26-MAR-2004) Clarke J.H., John Innes Centre, Colney Lane, Norwich, NR4 7UJ, UK

COMMENT AT denotes an activation tag dissociation transposon within a single line, ET an enhancer trap dissociation transposon, GT a gene trap dissociation transposon, MT a mis-expression enhancer trap dissociation transposon, SM a defective suppressor mutator transposon. _3 denotes a sequence derived from the 3' end of the transposon, _5 denotes a sequence derived from the 5' end of the transposon. BBSRC GARNET, ATIS project
 On-line seed stock requests: <http://nasc.nott.ac.uk/> NASC stock code: N165391.

FEATURES
 Location/Qualifiers
 1..530
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /db_xref="taxon:3702"
 /clone="AC005312"
 /ecotype="Landsberg erecta NASC stock code NW20"
 /note="Unknown"
 1..530
 /standard_name="GT_5.101721"

STS
 1..530
 /standard_name="GT_5.101721"

ORIGIN
 Query Match 17.2%; Score 447.8; DB 10; Length 530;
 Best Local Similarity 96.8%; Pred. No. 7.8e-64;
 Matches 489; Conservative 0; Mismatches 12; Indels 4; Gaps 3;

QY 956 GGTGTGAGCGCTTTAAATTTGAAATTAATTCCTTAGACAGTTTGAATTTTTC-TTTTTT 1014
 DB 28 GGTGTGAGCGCTTTAAATTTGAAATTAATTCCTTAGACAGTTTGAATTTTTC-TTTTTT 87

QY 1015 GACTAGTTTATTTATTTTGGAAATGATTCGATAGATCAAAAATACCTTGTCAATG 1074
 DB 88 GACTAGTTTATTTATTTTGGAAATGATTCGATAGATCAAAAATACCTTGTCAATG 147

QY 1075 GACTAAATGTCAGCGCGCTTTGCGCTTAAATCCA-GAAAAATGTTTCATGTCATATCGCT 1133
 DB 148 GACTAAATGTCAGCGCGCTTTGCGCTTAAATCCA-GAAAAATGTTTCATGTCATATCGCT 207

QY 1134 GAACCTTTTAAATTCGTAGACATGCGCCCATATGTTATAGTAGAATACATTAATAGATAGA 1193
 DB 208 GAACCTTTTAAATTCGTAGACATGCGCCCATATGTTATAGTAGAATACATTAATAGATAGA 267

QY 1194 TGCATACATATATATAACACACAAGATGATCACCTCGACATTCATATACCTTAATTCCT 1253
 DB 268 TGCATACAC--ATATATAACACACAAGATGATCACCTCGACATTCATATACCTTAATTCCT 325

QY 1254 GCAGACATAGTTAGTTTCTTACAAATTTATGACATGAATGTTCTGCTCTTCTCTCAC 1313
 DB 326 GCAGACATAGTTAGTTTCTTACAAATTTATGACATGAATGTTCTGCTCTTCTCTCAC 385

QY 1314 ATTAATTCATGTTCTTCTTATTTAAGTTTACCCAAATTTTGAATTAATTTGGCATATATG 1373
 DB 386 ATTAATTCATGTTCTTCTTATTTAAGTTTACCCAAATTTTGAATTAATTTGGCATATATA 445

QY 1374 AATTATACCAACATATTTATATGCGACATTTAAATCTATACGATGATTAACGGTTTAT 1433
 DB 446 AATTATACCAACATATCTATATGCGAACACTTAAATCTATACGAGGATTAACGGTTTAT 505

QY 1434 GGAGTAGACCGGAAAAATATATATGT 1458


```

Db      36184  TGTGTTTATATGTTTATGTTGATCTGTTTATGTAATCTTTATGAGCTGAGTGTGATCA 36243
QY      603  ATAGATATATGTTATGACGATATGCGCGGATAGAGAGAAAGAGTGGTACTTCTATGT 662
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      36244  ATATCTATGATGAGGTGCGAGCTTGGCAGCTATTTGGTGAAGAGGAGTGGTACTTCTATGT 36303
QY      663  GCCAAGAGATCGGAATATAGAAATGAGATAGACCGAAGCGAGTAACGACTTCAGATA 722
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      36304  GCCGAGAGATAGGAAGTATTCGCAACGCGGATCGTCCCAACCGTGTGACTACTTCTGGGTA 36363
QY      723  TTGGAAGCCACCGAGCTCATAGGATGATCAGATCGGAGACTTCTCGGCTATCCGATT 782
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      36364  TTGGAAGCAACCGAGCTGACAGGATGATTCGAACTGAAACTTCAGGTCAATGGGCT 36423
QY      783  AAGAAAAACCTAGTTTCTCTCTCTGTTAAAGCCCTTAAAGGCACCTGCTACTACTAGTTGGAT 842
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      36424  GAAGAAAAACCTAGTTTCTCTCTGTTAAAGCTCCCAAGGATCCGAACCACTTGGAT 36483
QY      843  CATGAACGAGTATCGTCTTCGACCATGAACCGAGTAACGAGTAACCAAGGTA 894
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      36484  TATGAATGATATCGTCTTCGCTCAGCATGAAGTGAACGATATCAAAAGGTA 36535

RESULT 10
AP007409
LOCUS
DEFINITION
  Lotus corniculatus var. japonicus chromosome 4 clone Lj146D13, ***
  SEQUENCING IN PROGRESS ***, 24 unordered pieces.
ACCESSION
  AP007409
VERSION
  HTG; HTGS PHASE1.
KEYWORDS
  Lotus corniculatus var. japonicus (Lotus japonicus)
SOURCE
  Lotus corniculatus var. japonicus
  ORGANISM
    Lotus corniculatus var. japonicus
      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
      Spermatophyta; Magnoliopsida; eudicotyledons; core eudicotyledons;
      rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
      Lotus.
REFERENCE
  1
  Kaneko,T., Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  Structural Analysis of a Lotus japonicus Genome. XI. Sequence
  Features and Mapping of Nine hundred twenty-one TAC Clones
  Unpublished
  2 (bases 1 to 91718)
  Sato,S.
  Direct Submission
  Submitted (26-OCT-2004) Shusei Sato, Kazusa DNA Research Institute,
  Department of Plant Gene Research; 2-6-7 Kazusa-Kamatari, Kisarazu,
  Chiba, 292-0818, Japan (E-mail: ssato@kazusa.or.jp,
  URL:http://www.kazusa.or.jp/, Tel:81-438-52-3935(ex.2337),
  Fax:81-438-52-3934)
  * NOTE: This is a 'working draft' sequence. It currently
  * consists of 24 contigs. The true order of the pieces
  * is not known and their order in this sequence record is
  * arbitrary. Gaps between the contigs are represented as
  * runs of N, but the exact sizes of the gaps are unknown.
  * This record will be updated with the finished sequence
  * as soon as it is available and the accession number will
  * be preserved.
  *
  * 1
  * 3414: contig of 3414 bp in length
  * 3415: gap of unknown length
  * 3515: contig of 916 bp in length
  * 4431: gap of unknown length
  * 4530: gap of unknown length
  * 4531: contig of 1486 bp in length
  * 6017: gap of unknown length
  * 6116: contig of 1428 bp in length
  * 6117: gap of unknown length
  * 7545: contig of 1003 bp in length
  * 7645: gap of unknown length
  * 8648: gap of unknown length
  * 8747: contig of 1288 bp in length
  * 10035: gap of unknown length
  * 10036: contig of 1773 bp in length
  * 11908: gap of unknown length
  * 12008: gap of unknown length
  * 12009: contig of 1489 bp in length

13498 13597: gap of unknown length
15598 15594: contig of 2057 bp in length
15655 15754: gap of unknown length
15755 18265: contig of 2511 bp in length
18266 18365: gap of unknown length
18366 20730: contig of 2365 bp in length
20731 20830: gap of unknown length
20831 22199: contig of 1269 bp in length
22199 22200: gap of unknown length
22200 25367: contig of 3168 bp in length
25368 29401: contig of 3834 bp in length
29402 32762: contig of 3361 bp in length
32763 3490: contig of 1328 bp in length
3491 34290: gap of unknown length
34291 38886: contig of 4596 bp in length
38887 38986: gap of unknown length
38987 43709: contig of 4723 bp in length
43710 43809: gap of unknown length
43810 47835: contig of 4026 bp in length
47836 47935: gap of unknown length
47936 54444: contig of 6509 bp in length
54445 54544: gap of unknown length
54545 60554: contig of 5910 bp in length
60555 70888: contig of 10334 bp in length
70889 70988: gap of unknown length
70989 88088: contig of 17100 bp in length
88089 88188: gap of unknown length
88189 91718: contig of 3530 bp in length.

FEATURES
      source
      1..91718
      /organism="Lotus corniculatus var. japonicus"
      /mol_type="genomic DNA"
      /variety="japonicus"
      /db_xref="taxon:34305"
      /chromosome="4"
      /clone="Lj146D13"
      /clone_lib="LJT library"
      /note="TAC clone:TM0525, synonym:Lotus japonicus"
      3415..3514
      /estimated_length=unknown
      4431..4530
      /estimated_length=unknown
      6017..6116
      /estimated_length=unknown
      7545..7644
      /estimated_length=unknown
      8648..8747
      /estimated_length=unknown
      10036..10135
      /estimated_length=unknown
      11909..12008
      /estimated_length=unknown
      13498..13597
      /estimated_length=unknown
      15655..15754
      /estimated_length=unknown
      18266..18365
      /estimated_length=unknown
      20731..20830
      /estimated_length=unknown
      22100..22199
      /estimated_length=unknown
      25368..25467
      /estimated_length=unknown
      29302..29401
      /estimated_length=unknown
      32763..32862
      /estimated_length=unknown
      34191..34290

```


AP008207_063 6300001 6410000
 AP008207_064 6400001 6510000
 AP008207_065 6500001 6610000
 AP008207_066 6600001 6710000
 AP008207_067 6700001 6810000
 AP008207_068 6800001 6910000
 AP008207_069 6900001 7010000
 AP008207_070 7000001 7110000
 AP008207_071 7100001 7210000
 AP008207_072 7200001 7310000
 AP008207_073 7300001 7410000
 AP008207_074 7400001 7510000
 AP008207_075 7500001 7610000
 AP008207_076 7600001 7710000
 AP008207_077 7700001 7810000
 AP008207_078 7800001 7910000
 AP008207_079 7900001 8010000
 AP008207_080 8000001 8110000
 AP008207_081 8100001 8210000
 AP008207_082 8200001 8310000
 AP008207_083 8300001 8410000
 AP008207_084 8400001 8510000
 AP008207_085 8500001 8610000
 AP008207_086 8600001 8710000
 AP008207_087 8700001 8810000
 AP008207_088 8800001 8910000
 AP008207_089 8900001 9010000
 AP008207_090 9000001 9110000
 AP008207_091 9100001 9210000
 AP008207_092 9200001 9310000
 AP008207_093 9300001 9410000
 AP008207_094 9400001 9510000
 AP008207_095 9500001 9610000
 AP008207_096 9600001 9710000
 AP008207_097 9700001 9810000
 AP008207_098 9800001 9910000
 AP008207_099 9900001 10010000
 AP008207_100 10000001 10110000
 AP008207_101 10100001 10210000
 AP008207_102 10200001 10310000
 AP008207_103 10300001 10410000
 AP008207_104 10400001 10510000
 AP008207_105 10500001 10610000
 AP008207_106 10600001 10710000
 AP008207_107 10700001 10810000
 AP008207_108 10800001 10910000
 AP008207_109 10900001 11010000
 AP008207_110 11000001 11110000
 AP008207_111 11100001 11210000
 AP008207_112 11200001 11310000
 AP008207_113 11300001 11410000
 AP008207_114 11400001 11510000
 AP008207_115 11500001 11610000
 AP008207_116 11600001 11710000
 AP008207_117 11700001 11810000
 AP008207_118 11800001 11910000
 AP008207_119 11900001 12010000
 AP008207_120 12000001 12110000
 AP008207_121 12100001 12210000
 AP008207_122 12200001 12310000
 AP008207_123 12300001 12410000
 AP008207_124 12400001 12510000
 AP008207_125 12500001 12610000
 AP008207_126 12600001 12710000
 AP008207_127 12700001 12810000
 AP008207_128 12800001 12910000
 AP008207_129 12900001 13010000
 AP008207_130 13000001 13110000
 AP008207_131 13100001 13210000
 AP008207_132 13200001 13310000
 AP008207_133 13300001 13410000
 AP008207_134 13400001 13510000
 AP008207_135 13500001 13610000

AP008207_136 13600001 13710000
 AP008207_137 13700001 13810000
 AP008207_138 13800001 13910000
 AP008207_139 13900001 14010000
 AP008207_140 14000001 14110000
 AP008207_141 14100001 14210000
 AP008207_142 14200001 14310000
 AP008207_143 14300001 14410000
 AP008207_144 14400001 14510000
 AP008207_145 14500001 14610000
 AP008207_146 14600001 14710000
 AP008207_147 14700001 14810000
 AP008207_148 14800001 14910000
 AP008207_149 14900001 15010000
 AP008207_150 15000001 15110000
 AP008207_151 15100001 15210000
 AP008207_152 15200001 15310000
 AP008207_153 15300001 15410000
 AP008207_154 15400001 15510000
 AP008207_155 15500001 15610000
 AP008207_156 15600001 15710000
 AP008207_157 15700001 15810000
 AP008207_158 15800001 15910000
 AP008207_159 15900001 16010000
 AP008207_160 16000001 16110000
 AP008207_161 16100001 16210000
 AP008207_162 16200001 16310000
 AP008207_163 16300001 16410000
 AP008207_164 16400001 16510000
 AP008207_165 16500001 16610000
 AP008207_166 16600001 16710000
 AP008207_167 16700001 16810000
 AP008207_168 16800001 16910000
 AP008207_169 16900001 17010000
 AP008207_170 17000001 17110000
 AP008207_171 17100001 17210000
 AP008207_172 17200001 17310000
 AP008207_173 17300001 17410000
 AP008207_174 17400001 17510000
 AP008207_175 17500001 17610000
 AP008207_176 17600001 17710000
 AP008207_177 17700001 17810000
 AP008207_178 17800001 17910000
 AP008207_179 17900001 18010000
 AP008207_180 18000001 18110000
 AP008207_181 18100001 18210000
 AP008207_182 18200001 18310000
 AP008207_183 18300001 18410000
 AP008207_184 18400001 18510000
 AP008207_185 18500001 18610000
 AP008207_186 18600001 18710000
 AP008207_187 18700001 18810000
 AP008207_188 18800001 18910000
 AP008207_189 18900001 19010000
 AP008207_190 19000001 19110000
 AP008207_191 19100001 19210000
 AP008207_192 19200001 19310000
 AP008207_193 19300001 19410000
 AP008207_194 19400001 19510000
 AP008207_195 19500001 19610000
 AP008207_196 19600001 19710000

Query Match 6.4%; Score 167.6; DB 15; Length 110000;
 Best Local Similarity 71.3%; Pred. No. 3.3e-18;
 Matches 221; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
 Qy 605 AGATATATGTATATGCGCGGCGATAGGAGAGAGAGTGGTACTTCTATGTGC 664
 Db 105436 AGTCGATGGGGTGCAGCAATGCGGCGCGATAGGAGAGAGTGGTCTTCTACGTGC 105495
 Qy 665 CAAGAGATCGGAATATAGAAATGGAGATAGACCGAACCGAGTACAGTATTT 724
 Db 105496 CTCGGGACAGAGTACAGACCGGCGCGGCGGCGGTACT 105555

Qy 725 GGAAGCCACCGAGCTGATAGATCGAGATCGAGACTTCTCGGCTATCGGATTAA 784
|||||
Db 105556 GGAAGCCACCGGCGCCGAGGATGATCCGAGCCGAGACAAACGCCCCATCGGCTCA 105615
Qy 785 AGAAACCTAGTTTCTTACTCTGTGTAAGCCCTAAAGGCACTCGTACTAGTTGGATCA 844
|||||
Db 105616 AGAAGACCTGTCTTCTTACTCCGCAAGGCCCCCAAGGGGTCGCGACGCTGGATCA 105675
Qy 845 TGAACGAGTATCGTCTTCCGACCATGAACCGAGAAGTACCAAGAGTATAAATCTTAC 904
|||||
Db 105676 TGAACGAGTATCGCTTCCCGCCGACCGGCTACCAAGGATCCCATCCATC 105735
Qy 905 TATAACTCTA 914
|||||
Db 105736 CATGAATCTA 105745

RESULT 12
AP008207_386
WPCOMMENT

Sequence split into 433 fragments LOCUS AP008207 Accession AP008207

| Fragment Name | Begin | End |
|---------------|----------|----------|
| AP008207_000 | 1 | 110000 |
| AP008207_001 | 100001 | 210000 |
| AP008207_002 | 200001 | 310000 |
| AP008207_003 | 300001 | 410000 |
| AP008207_004 | 400001 | 510000 |
| AP008207_005 | 500001 | 610000 |
| AP008207_006 | 600001 | 710000 |
| AP008207_007 | 700001 | 810000 |
| AP008207_008 | 800001 | 910000 |
| AP008207_009 | 900001 | 1010000 |
| AP008207_010 | 1000001 | 1110000 |
| AP008207_011 | 1100001 | 1210000 |
| AP008207_012 | 1200001 | 1310000 |
| AP008207_013 | 1300001 | 1410000 |
| AP008207_014 | 1400001 | 1510000 |
| AP008207_015 | 1500001 | 1610000 |
| AP008207_016 | 1600001 | 1710000 |
| AP008207_017 | 1700001 | 1810000 |
| AP008207_018 | 1800001 | 1910000 |
| AP008207_019 | 1900001 | 2010000 |
| AP008207_020 | 2000001 | 2110000 |
| AP008207_021 | 2100001 | 2210000 |
| AP008207_022 | 2200001 | 2310000 |
| AP008207_023 | 2300001 | 2410000 |
| AP008207_024 | 2400001 | 2510000 |
| AP008207_025 | 2500001 | 2610000 |
| AP008207_026 | 2600001 | 2710000 |
| AP008207_027 | 2700001 | 2810000 |
| AP008207_028 | 2800001 | 2910000 |
| AP008207_029 | 2900001 | 3010000 |
| AP008207_030 | 3000001 | 3110000 |
| AP008207_031 | 3100001 | 3210000 |
| AP008207_032 | 3200001 | 3310000 |
| AP008207_033 | 3300001 | 3410000 |
| AP008207_034 | 3400001 | 3510000 |
| AP008207_035 | 3500001 | 3610000 |
| AP008207_036 | 3600001 | 3710000 |
| AP008207_037 | 3700001 | 3810000 |
| AP008207_038 | 3800001 | 3910000 |
| AP008207_039 | 3900001 | 4010000 |
| AP008207_040 | 4000001 | 4110000 |
| AP008207_041 | 4100001 | 4210000 |
| AP008207_042 | 4200001 | 4310000 |
| AP008207_043 | 4300001 | 4410000 |
| AP008207_044 | 4400001 | 4510000 |
| AP008207_045 | 4500001 | 4610000 |
| AP008207_046 | 4600001 | 4710000 |
| AP008207_047 | 4700001 | 4810000 |
| AP008207_048 | 4800001 | 4910000 |
| AP008207_049 | 4900001 | 5010000 |
| AP008207_050 | 5000001 | 5110000 |
| AP008207_051 | 5100001 | 5210000 |
| AP008207_052 | 5200001 | 5310000 |
| AP008207_053 | 5300001 | 5410000 |
| AP008207_054 | 5400001 | 5510000 |
| AP008207_055 | 5500001 | 5610000 |
| AP008207_056 | 5600001 | 5710000 |
| AP008207_057 | 5700001 | 5810000 |
| AP008207_058 | 5800001 | 5910000 |
| AP008207_059 | 5900001 | 6010000 |
| AP008207_060 | 6000001 | 6110000 |
| AP008207_061 | 6100001 | 6210000 |
| AP008207_062 | 6200001 | 6310000 |
| AP008207_063 | 6300001 | 6410000 |
| AP008207_064 | 6400001 | 6510000 |
| AP008207_065 | 6500001 | 6610000 |
| AP008207_066 | 6600001 | 6710000 |
| AP008207_067 | 6700001 | 6810000 |
| AP008207_068 | 6800001 | 6910000 |
| AP008207_069 | 6900001 | 7010000 |
| AP008207_070 | 7000001 | 7110000 |
| AP008207_071 | 7100001 | 7210000 |
| AP008207_072 | 7200001 | 7310000 |
| AP008207_073 | 7300001 | 7410000 |
| AP008207_074 | 7400001 | 7510000 |
| AP008207_075 | 7500001 | 7610000 |
| AP008207_076 | 7600001 | 7710000 |
| AP008207_077 | 7700001 | 7810000 |
| AP008207_078 | 7800001 | 7910000 |
| AP008207_079 | 7900001 | 8010000 |
| AP008207_080 | 8000001 | 8110000 |
| AP008207_081 | 8100001 | 8210000 |
| AP008207_082 | 8200001 | 8310000 |
| AP008207_083 | 8300001 | 8410000 |
| AP008207_084 | 8400001 | 8510000 |
| AP008207_085 | 8500001 | 8610000 |
| AP008207_086 | 8600001 | 8710000 |
| AP008207_087 | 8700001 | 8810000 |
| AP008207_088 | 8800001 | 8910000 |
| AP008207_089 | 8900001 | 9010000 |
| AP008207_090 | 9000001 | 9110000 |
| AP008207_091 | 9100001 | 9210000 |
| AP008207_092 | 9200001 | 9310000 |
| AP008207_093 | 9300001 | 9410000 |
| AP008207_094 | 9400001 | 9510000 |
| AP008207_095 | 9500001 | 9610000 |
| AP008207_096 | 9600001 | 9710000 |
| AP008207_097 | 9700001 | 9810000 |
| AP008207_098 | 9800001 | 9910000 |
| AP008207_099 | 9900001 | 10010000 |
| AP008207_100 | 10000001 | 10110000 |
| AP008207_101 | 10100001 | 10210000 |
| AP008207_102 | 10200001 | 10310000 |
| AP008207_103 | 10300001 | 10410000 |
| AP008207_104 | 10400001 | 10510000 |
| AP008207_105 | 10500001 | 10610000 |
| AP008207_106 | 10600001 | 10710000 |
| AP008207_107 | 10700001 | 10810000 |
| AP008207_108 | 10800001 | 10910000 |
| AP008207_109 | 10900001 | 11010000 |
| AP008207_110 | 11000001 | 11110000 |
| AP008207_111 | 11100001 | 11210000 |
| AP008207_112 | 11200001 | 11310000 |
| AP008207_113 | 11300001 | 11410000 |
| AP008207_114 | 11400001 | 11510000 |
| AP008207_115 | 11500001 | 11610000 |
| AP008207_116 | 11600001 | 11710000 |
| AP008207_117 | 11700001 | 11810000 |
| AP008207_118 | 11800001 | 11910000 |
| AP008207_119 | 11900001 | 12010000 |
| AP008207_120 | 12000001 | 12110000 |
| AP008207_121 | 12100001 | 12210000 |
| AP008207_122 | 12200001 | 12310000 |

GENSCAN (<http://CCR-081.mit.edu/GENSCAN.html>), FGENESH (<http://www.softberry.com/>), GeneMark.hmm (<http://opal.biology.gatech.edu/GeneMark/>), GlimmerM (http://www.tigr.org/cdb/glimmer/glmr_form.html), RiceHMM (<http://rgp.dna.affrc.go.jp/RiceHMM/>), SplicePredictor (<http://bioinformatics.iastate.edu/cgi-bin/sp.cgi>), sim4 (<http://globin.cse.psu.edu/html/docs/sim4.html>), gap2 (<http://www.tigr.org/software/glimmer/>), BLASTN and BLASTX. The genomic sequence was searched against NCBI NonRedundant Protein database, nr (<ftp://ncbi.nlm.nih.gov/blast/db>) and the cDNA sequence database at RGP or DBJ. Protein homologies of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DBJ accession no.

A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence.

The orientation of the sequence is from T7 to SP6 of the PAC clone. This sequence of P0434C04 clone has an overlap with P0408G07 (DBJ: AP003379) clone at 5' end and an overlap with B109D03 (DBJ: AP003431) clone at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rgp.dna.affrc.go.jp/GenomeSeq.html>.

Location/Qualifiers

```

1. 150064
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clones="P0434C04"
2637..4109
/gene="P0434C04.1"
join(<2637..2646,3433..>4109)
/gene="P0434C04.1"
/notes="start and end point are not identified"
join(2637..2646,3433..4109)
/gene="P0434C04.1"
/codon_start=1
/product="AP2 domain-containing transcription factor-like protein"
/protein_id="BAD82116.1"
/db_xref="GI:56785228"
/translations="MPFVTLAFTPLANQIEGKSRTERDHGEVDLSSPEAMSRAE
CGGEEERCRVGRVRRRWCKVSEIRVPGTRERLWGSVATPEAAVAHDVAVFL
RGAGDGGGGATLNFERRAATYGGGAVALRSPRSVQSVASDAGMAADALVAARD
AAPAPATAPYDHCAGATTAKHDELARRGMYGAHAAGANARTSGERQLVCAEE
ISVDDMEILM"
complement(5213..8060)
/gene="P0434C04.2"
complement(join(5213..6148,6846..6973,7779..8060))
/gene="P0434C04.2"
/notes="supported by full-length cDNA(s): AK106349"
complement(join(5471..6148,6846..6973,7779..8037))
/gene="P0434C04.2"
/notes="contains EST(s): C74866(E60517)
contains full-length cDNA(s): AK106349
similar to Oryza sativa chromosome 5, P0426G01.14"
/codon_start=1
/product="unknown protein"
/protein_id="BAD82117.1"
/db_xref="GI:56785229"
/translations="MWALTNSGSSCLPARRTPPLAAGAGSLAAGPWCWSWR

```

```

ROPARWPKLAVSAGRSKSGRGGDEPKKNKAASSSSSGKGDASAPSGDVSNNE
LOSNDTVVPCNLISYWRDVRASPVVVKVQTVDAHTLPOAATDAPVHCLPRKWAHSIPM
PESCVVAABEELDNGCTFERTVILLRLGSRDAYDFPGVILNRPLTYKKHVNPSF
RNQATPFDSCSLFFGGVDMISFIMKRTTDDRPKIGFEVSGVCGFRTDEKASALL
KSGAVPEDLNIFYGVASYDYLSEIDQGYWHVTSCSGLISDLSLATDPSCLWTEI
LKLGGQYAEISQPKEDGS"
complement(12013..12168)
/gene="P0434C04.3"
complement(<12013..>12168)
/gene="P0434C04.3"
/notes="start and end point are not identified"
complement(12013..12168)
/gene="P0434C04.3"
/notes="predicted by GlimmerM etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD82118.1"
/db_xref="GI:56785230"
/translations="MEKEDPPPITHNALPDLWVHVHLMQYALPQSESPSHPSAE
SOVCLSN"
join(14818..14996,15784..15986,17059..17140,17333..17397,
17595..17694,17981..18022,18167..18208,18332..19026)
/gene="P0434C04.5"
join(14818..14996,15784..15986,17059..17140,17333..17397,
17595..17694,17981..18022,18167..18208,18332..19026)
/gene="P0434C04.5"
/notes="supported by full-length cDNA(s): AK070958"
complement(14954..15214)
/gene="P0434C04.4"
complement(14954..15214)
/gene="P0434C04.4"
/notes="hypothetical ORF
predicted by GlimmerM
this category is not included in IRGSP standard"
join(15802..15986,17059..17140,17333..17397,17595..17694,
17981..18022,18167..18208,18332..18613)
/gene="P0434C04.5"
/notes="contains full-length cDNA(s): AK070958"
/codon_start=1
/product="MADS-box protein RMADS207"
/protein_id="BAD82119.1"
/db_xref="GI:56785231"
/translations="MGRKIEIKRIENKTSQVTFCKRRNGLKKAYELAILCDAEIA
LIVFSRGLYFESNVNSTRTYKKSASTSGSAPVIDVNSHOYFOQAAKMRHO
IOTLQANPHLIGESIGNMTAKELSLNLEKIGIRSKKHELLSEIYMQKREA
DLQENNFLRAKVAERAEHDDQQAEDDEMAPAPAVGGSSSGTELEALPATFDR
EYTOPAPPYMLAAAAAAYSSDDHHQTHLHLYPKYKVDGSGKGLL"
21867..22061
/gene="P0434C04.6"
<21867..>22061
/gene="P0434C04.6"
/notes="start and end point are not identified"
21867..22061
/gene="P0434C04.6"
/notes="predicted by GENSCAN etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD82120.1"
/db_xref="GI:56785232"
/translations="MKEMVDLGGQRRPMIGTQICSAKYRPGEMDQDETEHQGHFCHP
DAAATIGVNAVVKRPPRRP"
join(23062..23674,24261..24346,25033..25178,25498..25614,
25709..25825,26539..26682,26780..27409)
/gene="P0434C04.7-1"
join(23062..23674,24261..24346,25033..25178,25498..25614,
25709..25825,26539..26682,26780..27407)
/gene="P0434C04.7-1"
/notes="supported by full-length cDNA(s): AK065093"
23069..24124
/gene="P0434C04.7-2"
join(<23069..23674,23798..>24124)
/gene="P0434C04.7-2"
/notes="supported by full-length cDNA(s): AK120493"

```

FEATURES
source


```

/gene="B1099D03.3"
/note="gag-pol polyprotein-like"
complement(join(5630..5986,6158..6445))
/gene="B1099D03.4"
complement(join(5630..5986,6158..6445))
/note="gag-pol polyprotein-like"
complement(6458..6835)
/gene="B1099D03.5"
complement(6458..6835)
/note="B1099D03.5"
complement(6458..6835)
/note="gag-pol polyprotein-like"
complement(9051..9561)
/gene="B1099D03.6"
complement(join(9051..9104,9202..9561))
/note="start and end point are not identified"
complement(join(9051..9104,9202..9561))
/gene="B1099D03.6"
/note="contains EST(s): AU032092(R3496)"
/codon_start=1
/product="unknown protein"
/protein_id="BAD82360.1"
/db_xref="GI:56784339"
/translation="MARALLRITRSCYPLSPRLPPPLHRRGAAPLRRFLSLSS
SAPSSAAATANSDDGKGGGGGDDHDLMSDLMGCQCEGTFKASGPGGQHRN
KRSGLRHRPTGIIAQLPERAVQCFTVASEF"
10369..10901
/gene="B1099D03.7"
join(<10369..10573,10879..>10901)
/gene="B1099D03.7"
/note="start and end point are not identified"
join(10369..10573,10879..10901)
/gene="B1099D03.7"
/note="predicted by FGENSEH etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD82361.1"
/db_xref="GI:56784340"
/translation="MAMSSSLFTIFSPLISLCSLCSAGSGGGRWRPAKGDAMA
MMAERSCARAVGKRRAGGQSGAHKNQ"
complement(11793..13156)
/gene="B1099D03.8"
complement(join(11793..11996,12814..12985,13083..>13156))
/note="B1099D03.8"
/note="start and end point are not identified"
complement(join(11793..11996,12814..12985,13083..13156))
/gene="B1099D03.8"
/note="predicted by FGENSEH etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD82362.1"
/db_xref="GI:56784341"
/translation="MYSLKRGKGGYKMKVAVLEFNKSVKYPKRGKATTPPIKLPLP
SLNKYKPKRKITPTPIYVIEISKEGDNNTNTVYSDIMQEQFAVARSIS
DTKLMEQNLCKKAIKSPSSNLRVALTRRIGLEEDRTILHS"
13741..14601
/gene="B1099D03.9"
13741..14601
/gene="B1099D03.9"
/note="probably inactive due to 5'exon missing in CDS
probably inactive due to including stop codon(s) in CDS
pseudogene, putative gag-pol polyprotein"
complement(17380..22093)
/gene="B1099D03.10"
complement(join(17380..17909,17981..18110,20523..20751,
21600..21632,21719..22093))
/gene="B1099D03.10"
/note="supported by full-length cDNA(s): AK071139"
complement(join(17634..17909,17981..18110,20523..20751,
21600..21632,21719..21869))
/gene="B1099D03.10"

/note="contains EST(s):
AU082754(S14363),AU030871(E560367),AU173000(E560367),D48242
(S14363)
contains full-length cDNA(s): AK071139,AK058384,AK058498"
/codon_start=1
/product="nucleic acid binding protein"
/protein_id="BAB92630.1"
/db_xref="GI:20804953"
/translation="MEGGGGGGGGGGGGGGGGGGGGGAGATATRTAEVFRDLRGRAGH
IKALTDDVERFYKLCDEPEKLNCLYGYNETWVTLPAEEVFPPIPAFGINPARDG
MNEKDMALVAHSDSWLLSVFYFGARFGRREARRLNMINNLTIFVVTGAAR
KQAEKTPNSSSKNSKPSKVSQKASRSKLSAPKDESGDDGEEEDDDHNTL
CGTGTNDGKDEFWICCNCEKWHGKVKITPARAEHRIKQKPCDCTNKARA"
complement(22303..22808)
/gene="B1099D03.11"
complement(join(<22303..22314,22581..>22808))
/note="B1099D03.11"
/note="start and end point are not identified"
complement(join(22303..22314,22581..22808))
/gene="B1099D03.11"
/note="predicted by FGENSEH etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD82363.1"
/db_xref="GI:56784342"
/translation="MAIADLLPLPASSSMSGLASGHWMLHVPFTVGASVFPFALIPK
TCVEIATGAPAPERQTICNESSSRRHLHCSLGOI"

Query Match 6.4% Score 167.6; DB 15; Length 191022;
Best Local Similarity 71.3% Pred. No. 2.9e-18;
Matches 221; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

Qy 605 AGATATATGATATGAGCTATGCGGCGATAGAGAGAAAGTGTCTTCTATGTGC 664
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
63409 AGTCGATGGGGTGCAGCAATGCGGCGATAGAGAGAAAGTGTCTTCTACGTGC 63468
Qy 665 CAAGAGATCGGAATATAGAAATGGAGATAGACCGACCGAGTAACGACTTCAGGATATT 724
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
63469 CTGGGACAGGAAGTACAGAACCGGGACCGGGCGAACCGGGTGACCGGGTACT 63528
Qy 725 GGAAGCCACCGGAGCTGATAGATGATCAGATCGGAGACTTCTCGGCGCTATCGGATTA 784
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
63529 GGAAGGACCGGGGCGCGCGATGATCCGAGCGAGAACACCGCCCATCGGCTCA 63588
Qy 785 AGAAACCCCTAGTTTCTACTCTGGTAAAGCCCTTAAGGCATCTGCTACTAGTTGGATCA 844
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
63589 AGAAGACGCTTGTCTTCTACTCTCGGCAAGGCCCGCCAGGCGCTCCGACGAGCTGGATCA 63648
Qy 845 TGAACGAGTATCGTCTTCCGCACCATCAACCGAGAGTACCAAGAGTATAATTTCTAC 904
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
63649 TGAACGAGTATCGCTTCCCGCCGACACCGCGCTACCAAGGTACCCATCCATC 63708
Qy 905 TATAACTCTA 914
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
63709 CATGAATCTA 63718

RESULT 15
AX654050
LOCUS AX654050 1212 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 3920 from Patent W003000898.
ACCESSION AX654050
VERSION AX654050.1 GI:29156864
KEYWORDS
SOURCE
ORYZA sativa
ORGANISM
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1
REFERENCE
AUTHORS Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,
Katagiri,F., Qian,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.
TITLE Plant genes involved in defense against pathogens

```

JOURNAL Patent: WO 03000898-A 3920 03-JAN-2003;
Syngenta Participations AG (CH)
FEATURES Location/Qualifiers
source 1. .1212
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"

ORIGIN

| | Query Match | 6.3% | Score 164; | DB 6; | Length 1212; |
|----|-----------------------|---|--------------------|-----------|--------------|
| | Best Local Similarity | 70.8%; | Pred. No. 3.7e-17; | | |
| | Matches 218; | Conservative 0; | Mismatches 90; | Indels 0; | Gaps 0; |
| QY | 605 | AGATATATGTATATGCGGCGGATAGGAGAGAAAGTGGTACTTCTATGTGC | 664 | | |
| DB | 161 | AGCTCGATGGGGTGCAGCAATGGCGCGATAGGGGAGAAGGAGTGGTTCTTCTACGTGC | 220 | | |
| QY | 665 | CAAGAGATCGGAATATAGAAATGGAGATAGACCGAACCGAGTAACGACTTCAGGATATT | 724 | | |
| DB | 221 | CTCGGACACAGGAAGTACAGAACCGGGACCGGCCGAACCGGGTGACGGCGTCGGGTACT | 280 | | |
| QY | 725 | GGAAAGCCACCGGAGCTGATAGGATGATCAGATCGGAGACTTCTCGGCCCTATCGGATTAA | 784 | | |
| DB | 281 | GGAGGCGACGGGGCGCGCCGATGATCCGAGCCCGAACAACCGCCCATCGGGCTCA | 340 | | |
| QY | 785 | AGAAACCCCTAGTTTCTACTCTGGTAAAGCCCTTAAAGGCACTCGTACTAGTTGGATCA | 844 | | |
| DB | 341 | AGAAGACGCTTGTCTTCTACTCCGGCAAGGCCCCCAAGGGCGTCCGACGAGCTGGATCA | 400 | | |
| QY | 845 | TGAACGAGTATCGTCTTCCGCACCATGAAACCGAGAAGTACCAAAAGGTATAAATCTTAC | 904 | | |
| DB | 401 | TGAACGAGTATCGCCTTCCCCCGCGACACCGCCGTACCAAGACTGAATCTCAC | 460 | | |
| QY | 905 | TATAACTC | 912 | | |
| DB | 461 | TATGCCGC | 468 | | |

Search completed: April 6, 2006, 12:58:34
Job time : 12948 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 6, 2006, 10:14:38 ; Search time 983 Seconds
(without alignments)
3188.295 Million cell updates/sec

Title: US-10-780-703-2

Perfect score: 2034

Sequence: 1 MATVSTTIIIPMSNQNNN.....MNPVDPGNRDHYTNIPFK 379

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 segs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame.p2n.model -DEV=xlh
-Q=/abes/ABSSWEB_spool/US10780703/runat_06042006_101348_6580/app_query.fasta_1
-DB=Published Applications NA Main -OFMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss07
-USER=US10780703 @CGN_1.1.1026 @runat_06042006_101348_6580 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|----------------------|
| 1 | 2034 | 100.0 | 1140 | 8 | US-10-780-703-1 |
| 2 | 2006.5 | 98.6 | 1444 | 6 | US-10-225-066A-115 |
| 3 | 2006.5 | 98.6 | 1444 | 7 | US-10-225-067-87 |
| 4 | 2006.5 | 98.6 | 1444 | 7 | US-10-374-780A-2475 |
| 5 | 2006.5 | 98.6 | 1444 | 9 | US-10-225-066A-115 |
| 6 | 1735 | 85.3 | 2606 | 8 | US-10-780-703-3 |
| 7 | 889 | 43.7 | 785 | 7 | US-10-424-599-121054 |

| | | | | | | |
|----|-------|------|------|---|----------------------|--------------------|
| 8 | 812 | 39.9 | 1433 | 7 | US-10-437-963-88434 | Sequence 88434, A |
| 9 | 781.5 | 38.4 | 1291 | 8 | US-10-425-115-10100 | Sequence 10100, A |
| 10 | 781 | 38.4 | 743 | 7 | US-10-437-963-81769 | Sequence 81769, A |
| 11 | 737 | 36.2 | 792 | 8 | US-10-425-115-173486 | Sequence 173486, A |
| 12 | 545 | 26.8 | 1321 | 7 | US-10-437-963-53771 | Sequence 53771, A |
| 13 | 535.5 | 26.3 | 1902 | 7 | US-10-424-599-100253 | Sequence 100253, A |
| 14 | 535 | 26.3 | 1644 | 7 | US-10-425-114-20520 | Sequence 20520, A |
| 15 | 527.5 | 25.9 | 1130 | 6 | US-10-225-066A-849 | Sequence 849, App |
| 16 | 527.5 | 25.9 | 1130 | 7 | US-10-374-780A-319 | Sequence 319, App |
| 17 | 527.5 | 25.9 | 1130 | 7 | US-10-225-066A-849 | Sequence 849, App |
| 18 | 526 | 25.9 | 390 | 8 | US-10-425-115-101230 | Sequence 101230, A |
| 19 | 525.5 | 25.8 | 1319 | 8 | US-10-425-115-121702 | Sequence 121702, A |
| 20 | 518 | 25.5 | 1257 | 6 | US-10-225-066A-287 | Sequence 287, App |
| 21 | 518 | 25.5 | 1257 | 7 | US-10-374-780A-2591 | Sequence 2591, App |
| 22 | 518 | 25.5 | 1257 | 9 | US-10-225-066A-287 | Sequence 287, App |
| 23 | 517.5 | 25.4 | 1152 | 7 | US-10-437-963-25359 | Sequence 25359, A |
| 24 | 517 | 25.4 | 1876 | 7 | US-10-374-780A-844 | Sequence 844, App |
| 25 | 516 | 25.4 | 1580 | 7 | US-10-425-114-20659 | Sequence 20659, A |
| 26 | 514 | 25.3 | 1032 | 7 | US-10-425-114-12858 | Sequence 12858, A |
| 27 | 509 | 25.0 | 740 | 7 | US-10-424-599-44069 | Sequence 44069, A |
| 28 | 509 | 25.0 | 1209 | 9 | US-10-737-164-6 | Sequence 6, Appli |
| 29 | 508.5 | 25.0 | 1125 | 6 | US-10-225-066A-807 | Sequence 807, App |
| 30 | 508.5 | 25.0 | 1125 | 7 | US-10-374-780A-97 | Sequence 97, Appl |
| 31 | 508.5 | 25.0 | 1125 | 9 | US-10-225-066A-807 | Sequence 807, App |
| 32 | 508 | 25.0 | 894 | 7 | US-10-374-780A-2607 | Sequence 2607, App |
| 33 | 508 | 25.0 | 894 | 7 | US-10-412-699B-879 | Sequence 879, App |
| 34 | 508 | 25.0 | 894 | 8 | US-10-495-918-145 | Sequence 145, App |
| 35 | 508 | 25.0 | 1260 | 6 | US-10-278-536-73 | Sequence 73, Appl |
| 36 | 508 | 25.0 | 1260 | 6 | US-10-225-066A-1035 | Sequence 1035, App |
| 37 | 508 | 25.0 | 1260 | 6 | US-10-302-267-177 | Sequence 177, App |
| 38 | 508 | 25.0 | 1260 | 9 | US-10-225-066A-1035 | Sequence 1035, App |
| 39 | 507.5 | 25.0 | 1886 | 7 | US-10-374-780A-838 | Sequence 838, App |
| 40 | 507 | 24.9 | 702 | 7 | US-10-437-963-2167 | Sequence 2167, App |
| 41 | 505 | 24.8 | 1851 | 7 | US-10-425-114-5337 | Sequence 5337, App |
| 42 | 505 | 24.8 | 1811 | 8 | US-10-425-115-137682 | Sequence 137682, A |
| 43 | 503 | 24.7 | 1352 | 7 | US-10-425-114-14719 | Sequence 14719, A |
| 44 | 503 | 24.7 | 1324 | 5 | US-10-286-264-29 | Sequence 29, Appl |
| 45 | 503 | 24.7 | 1324 | 6 | US-10-278-536-175 | Sequence 175, App |

ALIGNMENTS

RESULT 1
US-10-780-703-1
; Sequence 1, Application US/10780703
; Publication No. US20050034194A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Jong Seob
; APPLICANT: Kim, Yun Hee
; APPLICANT: Choi, Eun kyung
; APPLICANT: Yoo, So yeon
; APPLICANT: Ahn, Ji Hoon
; APPLICANT: Choi, Yang Do
; TITLE OF INVENTION: Gene Controlling Flowering Time of Plants and Method for
; TITLE OF INVENTION: Manipulating Flowering Time of Plant Using the Same
; FILE REFERENCE: 012679-105
; CURRENT APPLICATION NUMBER: US/10/780, 703
; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: KR 10-2003-10772
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1140
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1137)
; OTHER INFORMATION: cDNA of LOW1 gene
US-10-780-703-1
Alignment Scores:

QY 61 GluGluGluLeuIleGluPheTyrLeuArgArgLysValGluGlyLysArgPheAsnVal 80
DB 328 GAAGAGAACTCATAGAGTTTACCTTCGCGGAAAGTTGAAGCGAAACGCTTAAATGTA 387
QY 81 GluLeuIleThrPheLeuAspLeuTyrArgTyrAspProThrGluLeuProAlaMetAla 100
DB 388 GAACCTCATCTTTCCCTCGATCTTATCGCTATGATCCTTGGGAACCTTCTGCTATGGCG 447
QY 101 AlaIleGlyGluLysGluTyrPheTyrValProArgAspArgLysTyrArgAsnGly 120
DB 448 GCGATAGAGAGAAAGAGTGTGATCTTATGTGCGAAGAGATCGGAATATAGAAATGGA 507
QY 121 AspArgProAsnArgValThrThrSerGlyTyrTrpLysAlaThrGlyAlaAspArgMet 140
DB 508 GATAGACCGAACCGAGTAACGACTTCAGGATATTGGAAAGCCACCGGAGCTGATAGGATG 567
QY 141 IleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTyrSerGly 160
DB 568 ATCAGATCGGAGACTTCTCGGCTATCGGATTAAGAAACCCCTAGTTTCTACTCTGGT 627
QY 161 LysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeuProHisHis 180
DB 628 AAAGCCCTTAAGGACCTCGTACTAGTGGATCATGAACGAGTATCGTCTTCGCCACCAT 687
QY 181 GluThrGluLysTyrGlnLysAlaGluIleSerLeuCyArgValTyrLysArgProGly 200
DB 688 GAAACCGAAGTAGTACCAAGGCTGAATATCATTTGCGGAGTGTACAAAGGCCAGGA 747
QY 201 ValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsnHisAsnSer 220
DB 748 GTAGAGATCATCCATCGGTACCAAGTCTCTCTCCACAAGACATCATACCACTCA 807
QY 221 SerThrSerSerArgLeuAlaLeuArgGlnGlnHisHisSerSerSerSerAsnHis 240
DB 808 TCGACATCATCCGCTTTAGCGCTTAAGACAACAACACACCATTCATCTCTCTTAATCAT 867
QY 241 SerAspAsnAsnLeuAsnAsnAsnAsnLeuAsnLeuGluLysLeuSerThrGlu 260
DB 868 TCGACACAAACCTTAAACAACAACAACAACAATCTCGAAGAGCTCTCCACCGAA 927
QY 261 TyrSerGlyAspGlySerThrThrThrThrThrThrThrThrThrThrThrThrThr 280
DB 928 TATTCGGCGACGCGACGACCAACAACAACACCAACAGTAACTCTGACGTTTACCATT 987
QY 281 AlaLeuAlaAsnGlnAlaIleTyrArgProMetProTyrAspThrSerAsnAsnThrLeu 300
DB 988 GCTCTAGCCCAATCAAAACATATATCGTCCAATGCTTTACGACACAACGACAACACATTG 1047
QY 301 IleValSerThrArgAsnHisGlnAspAspGluThrAlaIleValAspAspLeuGln 320
DB 1048 ATAGTCTCTACGAGAAATCATACAGACGATGATGAATCTGCCATTGTTGAGATCTTCAA 1107
QY 321 ArgLeuValAsnTyrGlnIleSerAspGly----- 330
DB 1108 AGACTAGTTAACTTACCAAAATATCAGATGGAGGTAAACATCAATCACCATACTTTCAAATT 1167
QY 330 ----- 330
DB 1168 GCTCAACAGTTTTCATCATCTCAACAACAAATGCTTAACGCAACGCAATTAACATTTGGTG 1227
QY 331 -----AlaThrThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsn 345
DB 1228 GCTGGCGGCGACTACGCGCAACGCTAATGCTCAAACTCAAGCGCGGCTTAGCTATGAAC 1287
QY 346 MetIleProAlaGlyThrIleProAsnAsnAlaLeuTrpAspMetTrpAsnProIleVal 365
DB 1288 ATGATTCCTGCGAGGACGATTCCAACAATGCTTTGTTGGATATGTGATATCCAAATGTA 1347
QY 366 ProAspGlyAsnArgAspHisTyrThrAsnIleProPheLys 379
DB 1348 CCAGATGGAAACAGAGATCATCTATCTATCTATCTTTTAAAG 1389

RESULT 3
US-10-225-067-87
; Sequence 87, Application US/10225067
; Publication NO. US20040019925A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Creelman, Robert A.
; APPLICANT: Keddle, James
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Dubell, Arnold T.
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pineda, Omlair
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Broun, Pierre E.
; TITLE OF INVENTION: BIOCHEMISTRY-RELATED POLYNUCLEOTIDES AND
; FILE OF INVENTION: POLYPEPTIDES IN PLANTS
; FILE REFERENCE: 514442002042
; CURRENT APPLICATION NUMBER: US/10/225,067
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 87
; LENGTH: 1444
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (148)....(1392)
US-10-225-067-87
Alignment Scores:
Pred. No.: 1,37e-216 Length: 1444
Score: 2006.50 Matches: 379
Percent Similarity: 91.5% Conservative: 0
Best Local Similarity: 91.5% Mismatches: 0
Query Match: 98.6% Indels: 35
DB: Gaps: 1
US-10-780-703-2 (1-379) x US-10-225-067-87 (1-1444)
QY 1 MetAlaIleValSerSerThrThrSerIleIleProMetSerAsnGlnValAsnAsn 20
DB 148 ATGGCAATTGTATCTCTCCACAACAGCATCATTCCTCATGAGTAAACCAAGTCAACAATAAC 207
QY 21 GluLysGlyIleGluAspAsnAspHisArgGlyGlyGlnGluSerHisValGlnAsnGlu 40
DB 208 GAAAGAGGTATAGAGACAATGATCATAGAGCGCGCCCAAGAGAGTCAATGCCAAATGAA 267
QY 41 AspGluAlaAspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThr 60
DB 268 GATGAGCTGATGATCATGATCATGATCATGATCATGATCATGATCATGATCATGATCATGAT 327
QY 61 GluGluGluLeuIleGluPheTyrLeuArgArgLysValGluGlyLysArgPheAsnVal 80
DB 328 GAAGAGAACTCATAGAGTTTACCTTCGCGGAAAGTTGAAGCGCAACGCTTAAATGTA 387
QY 81 GluLeuIleThrPheLeuAspLeuTyrArgTyrAspProThrGluLeuProAlaMetAla 100
DB 388 GAACCTCATCTTCTCGATCTTATCGCTATGATCCTTGGGAACCTTCTGCTATGGCG 447
QY 101 AlaIleGlyGluLysGluTyrPheTyrValProArgAspArgLysTyrArgAsnGly 120

Db 448 GCGATAGGAGAGAAAGAGTGGTACTTCTATGTGCCAAGAGATCGGAAATATAGAAATGGA 507
 QY 121 AspArgProAsnArgValThrThrSerGlyTyrTyrPheAlaThrGlyAlaAspArgMet 140
 Db 508 GATAGACCGAAGCGAGTAACGACTTCAGGATATTTGAAAGCCACCGAGCTGATAGGATG 567
 QY 141 IleArgSerGluThrSerArgProIleGlyLeuLeuValPheTyrSerGly 160
 Db 568 ATCAGATCGAGACTTCTCGGCTATCGGATTAAGAAACCTAGTTTCTACTCTGT 627
 QY 161 LysAlaProLeuGlyThrArgThrSerTyrPheMetAsnGluTyrArgLeuProHisHis 180
 Db 628 AAAGCCCTAAAGGCACTCGTACTAGTTCGATCATGACGAGTATCGTCTCCGACCAT 687
 QY 181 GluThrGluGlyTyrGlnLysAlaGluIleSerLeuCysArgValTyrLysArgProGly 200
 Db 688 GAAACCGAGAAGTACCAAAAGGCTGAATATATCATTTGTCGAGTGTACAAAAGGCCAGGA 747
 QY 201 ValGluAspHisProSerValProArgSerLeuSerThrArgHisAsnHisAsnSer 220
 Db 748 GTAGAAGATCATCCATCGGTACACGCTTCTCTCCACAGACATCATTAACCATACTCA 807
 QY 221 SerThrSerSerArgLeuAlaLeuArgGlnGlnHisSerSerSerSerSerSerSer 240
 Db 808 TCGACATCATCCGTTTAGCTTTAGCAACAAACACACCATTCATCTCTCTCTAATCAT 867
 QY 241 SerAspAsnLeuAsnAsnAsnAsnAsnIleAsnAsnLeuGluLysLeuSerThrGlu 260
 Db 868 TCCGACAAACCTTTAAACAAACAAACAAACATCAACATCTCGAAGAGCTCTCCACCGAA 927
 QY 261 TyrSerGlyAspGlySerThrThrThrThrThrThrThrThrThrThrThrThrThr 280
 Db 928 TATTCGGCGAGCGGAGCAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 987
 QY 281 AlaLeuAlaAsnGlnAsnIleTyrArgProMetProTyrAspThrSerAsnAsnThrLeu 300
 Db 988 GCTCTAGCCAATCAAAACATATATCTGCAATGCTTACGACAAACAAACACATTTG 1047
 QY 301 IleValSerThrArgAsnHisGlnAspAspAspGluThrAlaIleValAspAspLeuGln 320
 Db 1048 ATAGTCTCTACGAGAAATCATCAAGACGATGATGAACCTGCCATTTGTCGACATCTTCAA 1107
 QY 321 ArgLeuValAsnTyrGlnIleSerAspGly----- 330
 Db 1108 AGACTAGTTAACTTACCAAAATATCAGATGAGGTAAATCAATCAACCAATCTTTCAAAT 1167
 QY 330 ----- 330
 Db 1168 GCTCAACAGTTTCATCATCTACTCAACAAACAAATGCTTAACGCAACGCAATTAATGTTG 1227
 QY 331 -----AlaThrThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsn 345
 Db 1228 GGTGCGCGACTACAGCGAACCGCTTAATGCTTCAAACTCAAGCGCGGTAGCTATGAAC 1287
 QY 346 MetIleProAlaGlyThrIleProAsnAsnAlaLeuTyrAspMetTyrAsnProIleVal 365
 Db 1288 ATGATTTCTCGAGAAACGATTCCAACAAATGCTTTGTGGATGATGTTGGAATCCATAGTA 1347
 QY 366 ProAspGlyAsnArgAspHisTyrThrAsnIleProPheLys 379
 Db 1348 CCAGATGGAACACAGATCACTATCACTAATATTTCTTTTAAAG 1389

RESULT 4

US-10-374-780A-2475
 ; Sequence 2475, Application US/10374780A
 ; Publication No. US20040019927A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sherman, Bradley K
 ; APPLICANT: Kiechmann, Jose Luis
 ; APPLICANT: Jiang, Cai-Zhong
 ; APPLICANT: Heard, Jacqueline E
 ; APPLICANT: Haake, Volker
 ; APPLICANT: Creelman, Robert A

; APPLICANT: Ratcliffe, Oliver
 ; APPLICANT: Adam, Luc J
 ; APPLICANT: Reuber, T. Lynne
 ; APPLICANT: Keddier, James E
 ; APPLICANT: Brown, Pierre E
 ; APPLICANT: Pilgrim, Marsha L
 ; APPLICANT: Dubell III, Arnold T
 ; APPLICANT: Pineda, Onaira
 ; APPLICANT: Xu, Guo-Liang
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
 ; FILE REFERENCE: MBI-0047 CIP
 ; CURRENT APPLICATION NUMBER: US/10/374,780A
 ; CURRENT FILING DATE: 2003-02-25
 ; PRIOR APPLICATION NUMBER: 09/837,944
 ; PRIOR FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: 60/310,847
 ; PRIOR FILING DATE: 2001-08-09
 ; PRIOR APPLICATION NUMBER: 09/934,455
 ; PRIOR FILING DATE: 2001-08-22
 ; PRIOR APPLICATION NUMBER: 60/336,049
 ; PRIOR FILING DATE: 2001-11-19
 ; PRIOR APPLICATION NUMBER: 60/338,692
 ; PRIOR FILING DATE: 2001-12-11
 ; PRIOR APPLICATION NUMBER: 10/171,468
 ; PRIOR FILING DATE: 2002-06-14
 ; PRIOR APPLICATION NUMBER: 10/225,066
 ; PRIOR FILING DATE: 2002-08-09
 ; PRIOR APPLICATION NUMBER: 10/225,067
 ; PRIOR FILING DATE: 2002-08-09
 ; PRIOR APPLICATION NUMBER: 10/225,068
 ; PRIOR FILING DATE: 2002-08-09
 ; NUMBER OF SEQ ID NOS: 2906
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 2475
 ; LENGTH: 1444
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; OTHER INFORMATION: G962
 US-10-374-780A-2475
 Alignment Scores:
 Pred. No.: 1,378-216 Length: 1444
 Score: 2006.50 Matches: 379
 Percent Similarity: 91.5% Conservative: 0
 Best Local Similarity: 91.5% Mismatches: 0
 Query Match: 98.6% Indels: 35
 DB: 7 Gaps: 1
 US-10-780-703-2 (1-379) x US-10-374-780A-2475 (1-1444)
 QY 1 MetAlaIleValSerSerThrThrSerIlelleProMetSerAsnGlnValAsnAsn 20
 Db 148 ATGGCAATTTGATCTCCACCAACCAAGCATATCCCATGAGTAACCAAGTCAACATAAC 207
 QY 21 GluLysGlyIleGluAspAsnAspHisArgGlyGlyGlnGluSerHisValGlnAsnGlu 40
 Db 208 GAAAGAGTATAGAACACATGATCTAGAGCGGCCCAAGAGAGTCATGTCCTCAATATGA 267
 QY 41 AspGluAlaAspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThr 60
 Db 268 GATGAAGCTGATGATCATGATCATGATCATGATGATGATGATGATGATGATGATGATGAT 327
 QY 61 GluGluGluLeuIleGluPheTyrLeuArgArgLysValGluGlyLysArgPheAsnVal 80
 Db 328 GAAGAAGAACTCATAGAGTTTTTACCTTCGCCGAAAGTTGAAGGCAACCGCTTTAATGA 387
 QY 81 GluLeuIleThrPheLeuAspLeuTyrArgTyrAspProTyrGluLeuProAlaMetAla 100
 Db 388 GAACATCATCATCTTCTCGATCTTTATCGCTATGATCTTCTGGAACTTCTCTCTATGGCG 447
 QY 101 AlaIleGlyGluLysGluTyrTyrPheTyrValProArgAspArgLysTyrArgAsnGly 120

448 CGCATAGGAGAGAAAGAGTGGTACTTCTATGTGCCAAGAGATCGGAAATATAGAAATGGA 507
121 AspArgProAsnArgValThrThrSerGlyTyrTrpLysAlaThrGlyValAspArgMet 140
508 GATAGACCGAACCGAGTACGACTTCAGATATTGGAAAGCCACCGGAGCTGATAGGATG 567
141 IleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTyrSerGly 160
568 ATCAGATCGGAGACTTCTCGGCTATCGGCTATCGGCTATCGGCTATCGGCTATCGGCT 627
161 LysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeuProHisHis 180
628 AAAGCCCTAAAGGCACCTCGTACTAGTTGGATCATGAACGAGTATCGTCTCCGACCAT 687
181 GluThrGluLysTyrGlnLysAlaGluLeuSerLeuCysArgValTyrLysArgProGly 200
688 GAAACCGAGAGTACCAAAAGGCTGAATATCATTTGTCGAGGTGTACAAAGGCCAGGA 747
201 ValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsnHisAsnSer 220
748 GTAGAAGATCATCCATCGTACCAGTTCTCTCTCCCAAGACATCATACCACTCA 807
221 SerThrSerArgLeuAlaLeuArgGlnGlnHisHisSerSerSerSerSerAsnHis 240
808 TCGACATCATCCCGTTAGCCCTTAAGCAACAACAACACCATTCATCTCTCTTAATCAT 867
241 SerAspAsnLeuAsnAsnAsnAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 260
868 TCCGACCAACACCTTAAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 927
261 TyrSerGlyAspGlySerThrThrThrThrThrThrThrThrThrThrThrThrThr 280
928 TATTCCGGCGAGCGGACGACACCAACACGACCAACACGATTAACCTCGAGTTACCAT 987
281 AlaLeuAlaAsnGlnAsnIleTyrArgProMetProTyrAspThrSerAsnAsnThrLeu 300
988 GCTCTAGCCAAATCAAAACATATATCTGTCCTCAATGCTTACGACACAAAGCAACA 1047
301 IleValSerThrArgAsnHisGlnAspAspGluThrAlaIleValAspAspLeuGln 320
1048 ATAGTCTCTACGAAATCATCAAGACGATGATGAACCTGCCATTTGTGAGCATCTTCAA 1107
321 ArgLeuValAsnTyrGlnIleSerAspGly 330
1108 AGACTAGTTAACTACCAATATCATGAGTGGAGGTAACTCAATCACTTTCAAT 1167
330 ----- 330
1168 GCTCAACAGTTTTCATCATCTCAACCAACAAATGCTTAACGCAACGCAATTTGGTG 1227
331 -----AlaThrThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsn 345
1228 GCTCGGCGGACTACGCGACACGCTAATGCTCAAACTCAAGCGGCGTTAGCTATGAAC 1287
346 MetIleProAlaGlyThrIleProAsnAsnAlaLeuTyrAspMetTyrAsnProIleVal 365
1288 ATGATTCTCGAGAACGATTCGAACCAATGCTTTGGGATATGTGAATTCATATGTA 1347
366 ProAspGlyAsnArgAspHisTyrThrAsnIleProPheLys 379
1348 CCAGATGGAACACAGATCACTATATCTTTTAAAG 1389

RESULT 5

US-10-225-066A-115
; Sequence 115, Application US/10225066A
; Publication No. US20050160493A9
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBELL, Arnold T
; APPLICANT: HEARD, Jacqueline E

APPLICANT: PILGRIM, Marsha L
APPLICANT: JIANG, Cai-Zhong
APPLICANT: REUBER, T. Lynne
APPLICANT: CREELMAN, Robert A
APPLICANT: PINEDA, Omaira
APPLICANT: YU, Guo-Liang
APPLICANT: BROUN, Pierre E
TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
FILE REFERENCE: MBI0036-2 US
CURRENT APPLICATION NUMBER: US/10/225,066A
CURRENT FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: 09/837,444
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/310,847
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 60/336,049
PRIOR FILING DATE: 2001-12-05
PRIOR APPLICATION NUMBER: 60/338,692
PRIOR FILING DATE: 2001-12-11
PRIOR APPLICATION NUMBER: 10/171,468
PRIOR FILING DATE: 2002-06-14
NUMBER OF SEQ ID NOS: 1122
SOFTWARE: PatentIn version 3.1
SEQ ID NO 115
LENGTH: 1444
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-10-225-066A-115

Alignment Scores:
Pred. No.: 1,37e-216 Length: 1444
Score: 2006.50 Matches: 379
Percent Similarity: 91.5% Conservative: 0
Best Local Similarity: 91.5% Mismatches: 35
Query Match: 98.6% Indels: 1
DB: 9 Gaps: 1
US-10-780-703-2 (1-379) x US-10-225-066A-115 (1-1444)
Qy 1 MetAlaIleValSerSerThrThrSerIleIleProMetSerAsnGlnValAsnAsnAsn 20
Db 148 ATGGCAATTGTATCTCTCCCAACAAGCATCATTTCCCATGAGTAACCAAGTCAACAATAAC 207
Qy 21 GluLysGlyIleGluAspAsnAspHisArgGlyGlnGluSerHisValGlnAsnGlu 40
Db 208 GAAAAAGTATAGAAACAATGATCATAGAGCGGCCCAAGAGATCATGTCCAAATGAA 267
Qy 41 AspGluAlaAspAspHisAspMetValMetProGlyPheArgPheHisProThr 60
Db 268 GATGAAGTCTGATCATGATCATGACATGTCATGCCGATTTAGATTCCATCTTACC 327
Qy 61 GluGluIleuLeuGluPheTyrLeuArgLysValGluGlyLysArgPheAsnVal 80
Db 328 GAAGAAGAACTCATAGAGTTTACCTTCGCGGAAAAAGTTGAAGGCAACGCTTTAATGTA 387
Qy 81 GluLeuIleThrPheLeuAspLeuTyrArgTyrAspProTyrGluLeuProAlaMetAla 100
Db 388 GAATCATCATCTTCTCTCGATCTTTATCGCTATGATCCTTGGGAACCTTCTGCTATGCG 447
Qy 101 AlaIleGlyLysGluTyrTrpTyrPheTyrValProArgAspArgLysTyrArgAsnGly 120
Db 448 GCGATAGAGAGAAAGAGTGTACTTCTATGTGCCAAGAGATCGAATAATATAGAAATGGA 507
Qy 121 AspArgProAsnArgValThrThrSerGlyTyrTrpLysAlaThrGlyAlaAspArgMet 140
Db 508 GATAGACCGAACCGAGTAAACGACTTCAGGATATTGGAAAGCCACCGGAGCTGATAGGATG 567
Qy 141 IleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTyrSerGly 160
Db 568 ATCAGATCGGAGACTTCTCGGCTATCGGCTATCGGCTATCGGCTATCGGCTATCGGCT 627
Qy 161 LysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeuProHisHis 180


```
Db 628 AAAGCCCTAAAGGCACTCGTACTAGTTGGATCATGAACGAGTATCGTCTTCGCGACCAT 687
Qy 181 GluThrGluLysTyrGlnLysAlaGluIleSerLeuCyshArgValTyrLysArgProGly 200
Db 688 GAAACCGAGAGTACCAAAAGGCTGAAATATCATTTGCGCGAGTGTACAAAAGGCCAGGA 747
Qy 201 ValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsnSer 220
Db 748 GTAGAGATCATCATCGGTACCAAGTCTCTCTCCACAGACATCATAACTAATCTCA 807
Qy 221 SerThrSerSerArgLeuAlaLeuArgGlnGlnGlnHisHisSerSerSerAsnHis 240
Db 808 TCACATCATCCCGTTAGCGTTAAGACAAACAACACCATTCATCTCTCTAATCAT 867
Qy 241 SerAspAsnAsnLeuAsnAsnAsnAsnIleAsnAsnLeuGluLysLeuSerThrGlu 260
Db 868 TCCGACAAACCACTTAACAACAACAACATCAACATCTCAGAGAGCTCTCCACCGAA 927
Qy 261 TyrSerGlyAspGlySerThrThrThrThrThrThrAsnSerAsnSerAspValThrIle 280
Db 928 TATTCGGCGACGCGACACACACACACACACACACACACACACACACACACATTG 987
Qy 281 AlaLeuAlaAsnGlnAsnIleTyrArgProMetProTyrAspThrSerAsnAsnThrLeu 300
Db 988 GCTTAGCCCAATCAAAACATATATCGTCCCAATGCTTACGACACAAGCAACACACATTG 1047
Qy 301 IleValSerThrArgAsnHisGlnAspAspAspGluThrAlaIleValAspAspLeuGln 320
Db 1048 ATAGTCTCTACGAGAATCATCAAGACGATGATGAACTGCCATTGTTGACGATCTTCAA 1107
Qy 321 ArgLeuValAsnTyrGlnIleSerAspGly- 330
Db 1108 AGACTAGTTAACTACCAATATCAGATGGAGGTAAACATCAATCACCACATCTTTCAAAT 1167
Qy 330 ----- 330
Db 1168 GCTCAACAGTTTCATCATCTCAACAACAATAATGCTAAAGCAACGCAATTACAAATTTGGTG 1227
Qy 331 -----AlaThrThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsn 345
Db 1228 GCTCGCGCGACTACGCGACACAGCTTAATGCTCAAACTCAAGCGCGTTAGCTATGNAAC 1287
Qy 346 MetIleProAlaGlyThrIleProAsnAsnAlaLeuTyrAspMetTyrAsnProIleVal 365
Db 1288 ATGATTCCTGCAGAACGATTCACAAACATGCTTTGTGGATATGGAATCCCAATAGTA 1347
Qy 366 ProAspGlyAsnArgAspHisTyrThrAsnIleProPheLys 379
Db 1348 CCAGATGGAACACAGATCACTATCTATCTATCTCTTTTAAAG 1389

RESULT 6
US-10-780-703-3
; Sequence 3, Application US/10780703
; Publication No. US20050034194A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Jong Seob
; APPLICANT: Kim, Yun Hee
; APPLICANT: Choi, Eun Kyung
; APPLICANT: Yoo, So Yeon
; APPLICANT: Ahn, Ji Hoon
; APPLICANT: Choi, Yang Do
; TITLE OF INVENTION: Gene Controlling Flowering Time of Plants and Method for
; FILE REFERENCE: 012679-105
; CURRENT APPLICATION NUMBER: US/10780, 703
; PRIOR FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: KR 10-2003-10772
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2606
; TYPE: DNA
```

```
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)....(2606)
; OTHER INFORMATION: genomic DNA of LOV1 gene
US-10-780-703-3

Alignment Scores:
Pred. No.: 1-96e-185 Length: 2606
Score: 1735.00 Matches: 378
Percent Similarity: 43.5% Conservative: 0
Best Local Similarity: 43.5% Mismatches: 1
Query Match: 85.3% Indels: 1
DB: 8 Gaps: 3

US-10-780-703-2 (1-379) x US-10-780-703-3 (1-2606)
Qy 1 MetAlaIleValSerSerThrThrSerIleIleProMetSerAsnGlnValAsnAsn 20
Db 1 ATGGCAATTTGTATCTCTCCACAAAGCATCATTCCTCATGAGTAACCAAGTCAACAATAAC 60
Qy 21 GluLysGlyIleGluAspAsnAspHisArgGlyGlyGlnGluSerHisValGlnAsnGlu 40
Db 61 GAAAAAGGTATAGAAAGACAATCATAGAGCGGCCAAGAGAGTCATGTCCTCAAAATGAA 120
Qy 41 AspGluAlaAspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThr 60
Db 121 GATGAAGCTGATGATCATGATCATGATGTCATGTCATGTCGCGGATTTAGATTCCTATCCTACC 180
Qy 61 GluGluGluLeuIleGluPheTyrLeuArgArgLysValGluGlyLysArgPheAsnVal 80
Db 181 GAAGAAGAACTCATAGAGTTTTACCTTCGCCGAAAGTTGAAGGCAAAAGCTTTAATGTA 240
Qy 81 GluLeuIleThrPheLeuAspLeuTyrArgTyrAspProTyrGluLeuPro- 97
Db 241 GAACATCATCACTTCTCGATCTTTATCGTATGATCCTTGGAACTTCTCGTAAATAT 300
Qy 97 ----- 97
Db 301 ACATTACATAAACAACACATAAATCATCTCAAACTATTGGAATCTTAATTTCTATTCA 360
Qy 97 ----- 97
Db 361 TATGTTAAGATCTTCTCTCTTATCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
Qy 97 ----- 97
Db 421 ATATATGTACCTACCTCCTTATGAAGTATTAATATGTGTGATCGATTAACAATTTCTCAATAT 480
Qy 97 ----- 97
Db 481 CTTTAAACGCTTCTCCCTCTTTAGTTCTTTCTTAAATTAACCTAATTAACAACCTACA 540
Qy 97 ----- 97
Db 541 TATATATCATAGATATACAAATATGTGTATGTTTTCATAATTAGCTTATGTATGTTAA 600
Qy 98 -----AlaMetAlaAlaIleGlyGluLysGluTyrPheTyr 110
Db 601 TCATAGATATATGATATGTCAGCTATGCGCGCATAGGAGAGAAAGAGTGGTACTTCTAT 660
Qy 111 ValProArgAspArgLysTyrArgAsnGlyAspAlqProAsnArgValThrThrSerGly 130
Db 661 GTGCCAAGAGATCGGAAATATAGAAATGAGATAGACCGAACCGAGTACGACTTCAGGA 720
Qy 131 TyrTyrLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGly 150
Db 721 TATTGGAAGCCACCGAGCTGATAGATGATCAGATCGAGACTTCTCGGCTATCGGA 780
Qy 151 LeuLysLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTyr 170
Db 781 TTAAGAAAAACCTAGTTTCTTACTCTGTGTAAGGCCCTTAAAGGCACTCGTACTAGTTGG 840
```

171 IleMetAsnGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLysAla- 188
 841 ATCATGAACGAGTATCGTCTCCGACCATGAACCGAGAGTACCAAAAGGT-ATAAAT 899
 188 ----- 188
 900 TCTACTAATCTATATATATATCTATTTATCATACATACATAGATATAACCCCTAGCTAGGTG 959
 188 ----- 188
 960 GTGAGGCTTTAAATTTGAATTAATCCCTAGACAGTTTGAATTTTCTTTTGTGACTA 1019
 188 ----- 188
 1020 GTTTTATTTATTTTGAATTTGATTCGATAAGATCAAAAATFACTTGTGAATGGACTA 1079
 188 ----- 188
 1080 AATGTCAGCGCGCTTTGCGCTTAAATCCAGAAAAATGTTTCATGTCATATGCGTGAATC 1139
 188 ----- 188
 1140 TTTAAATTTGTAGACATGGCCCATATGTTATAGTAGAATACATTAATAGATAGATGCA 1199
 188 ----- 188
 1200 CACATATATAAACAACAAGTATCACATCGACATTCATATACCTTAATTCGACAGAG 1259
 188 ----- 188
 1260 ACATAGTTAGTTTCTTACAATTTATGACATGAATGTTCTGCTCTTCTCTCACATTAAT 1319
 188 ----- 188
 1320 TCATGCTCTTATTTAAGTTTACCAACATTTTGTGAATAAATTTGGCATATATGAATTA 1379
 188 ----- 188
 1380 ACCAACATATTTATATCGGAACATTTAAATCTATACGAATGATAACGGTTTATGGAGTA 1439
 188 ----- 188
 1440 GACCGAAAAAATATATGATATACGGAATAATGACAAATGATAGATAAATACATTTTGTGG 1499
 188 ----- 188
 1500 CTCTTCGACTTATATGTCGTCAACATTTGAACCAATAAATTTATAAAATTTTCTATGTA 1559
 188 ----- 188
 1560 TATATATGATATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1619
 188 ----- 188
 1620 CTATGCTATGATTTGACATTCATCTGTTACTAATAAATAAATAAATAAATAAATAAATA 1679
 188 ----- 188
 1680 TATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1739
 188 ----- 188
 1740 TACTCTAAATTTCTTAAACGGTCTATCTTTTCTGCTAATGCTAATCTTACATAGTTT 1799
 188 ----- 188
 1800 GTGAATCTTCTTCAAAACCATATCTTCGATAAATGATATTTTTCATAGATATTTGTTAGT 1859
 188 ----- 188
 1860 CTATATTTGATAATTTGATATATGATCAAGTCTCTAATCAATGCTCATGCTATATAATTA 1919
 189 -----GluIleSerLeuCysArgValTyrLysArgProGlyValGluAspHisProSer 206

1920 TAGGCTGAATATATCATTTGTGCCAGTGTACAAAAGGCCAGAGTAGAAGATCATCCATCG 1979
 207 ValProArgSerLeuSerThrArgHisHisAsnHisAsnSerSerThrSerSerArgLeu 226
 1980 GTACCAAGTCTCTCTCCACAGACATCATACCATTAACATCATCATCATCCGCTTTA 2039
 227 AlaLeuArgGlnGlnHisHisSerSerSerSerSerSerSerSerSerSerSerSerSer 246
 2040 GCCTTAAGACAAACACACACACCATTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2099
 247 AsnAsnAsnAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 266
 2100 AACAAACAAACATCAACATCTCGAAGCTCTCCACCAATATTCGGCGAGCGAGC 2159
 267 Thr 286
 2160 ACAACAAACACGACCAACACAGTAACCTCTGACGTTACCAATTCCTAGCCAAATCAAAAC 2219
 287 IleTyrArgProMetProTyrAspThrSerAsnAsnThrLeuIleValSerThrArgAsn 306
 2220 ATATATCGTCCATGCTTACGACACACACACACACACATTTGATGCTCTCTACGAGAAAT 2279
 307 HisGlnAspAspAspGluThrAlaIleValAspAspLeuGlnArgLeuValAlaSerGln 326
 2280 CATCAAGACGATGATGAACCTGCACTGTTGACGATCTTCAAAGACTAGTTAACTACCAA 2339
 327 IleSerAspGly----- 330
 2340 ATATCAGATGGAGGTAACATCAATCAACATATCTTCAAATTCGTCAACAGTTTTCATCAT 2399
 331 -----Ala 331
 2400 ACTCAACAAACAAATGCTAACGCAACGCAATTAACATTTGGTGGTGGCGGACATACAGCG 2459
 332 ThrThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsnMetIleProAlaGlyThr 351
 2460 ACAACGCTAATGCTCAAACTCAAGCGGCTTAGCTATGAACATGATTTCTCTGCGAGAACG 2519
 352 IleProAsnAsnAlaLeuTyrAspMetTyrAsnProIleValProAspGlyAsnArgAsp 371
 2520 ATTCCAAACAAATGCTTTGTTGGATATGTGGAATTCCTCAATAGTACCAGATGGAACAGAGAT 2579
 372 HisTyrThrAsnIleProPheLys 379
 2580 CACTATATAATATTCCTTTTAAG 2603

RESULT 7
 US-10-424-599-121054
 ; Sequence 121054, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 121054
 ; LENGTH: 785
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_80317C.1
 US-10-424-599-121054

Alignment Scores: 2.63e-90 Length: 785
 Pred. No.: 889.00 Matches: 171
 Score:

Percent Similarity: 77.2% Conservative: 12
Best Local Similarity: 72.2% Mismatches: 16
Query Match: 43.7% Indels: 38
DB: 7 Gaps: 3

US-10-780-703-2 (1-379) x US-10-424-599-121054 (1-785)

```
QY 35 SerHisValGluAsnGluAspGluAla----- 43
DB 128 AGCCATAGCCAGACGAGGATGCCACCACCGCCTCCACCACCAAGCCCTCGAC 187
QY 44 -----AspAspHisAspHisAspMetValMetProGlyPheArgPheHisPro 59
DB 188 GTCGACGATATGATGATCAGGACGACATGCTATGCCGCTTCGCTTCACCCCA 247
QY 60 ThrGluGluGluLeuLeuGluPheTyLeuArgArgLysValGluGlyLysArgPheAsn 79
DB 248 ACTGAAGAAGAGCTCGTGGAAATCTACCTTCGCCGTAAAGTGAGGAAAGCGTTTCAAC 307
QY 80 ValGluLeuLeuThrPheLeuAspLeuTyArgTyArgTyArgProTropGluLeuProAlaMet 99
DB 308 GTTGAGCTATTACTTTCCTCGATCTTTATCGCTATGACCTTGGGAGCTTCCGCCCTTG 367
QY 100 AlaAlaIleGlyGluLysGluTrpTyPheTyValProArgAspArgLysTyArgAsn 119
DB 368 GCAGCTATTGGTGAGAGGAGTGACTTCTATGTGCCAGAGATAGGAAGTATCGCAAC 427
QY 120 GlyAspArgProAsnArgValThrThrSerGlyTyTrpLysAlaThrGlyAlaAspArg 139
DB 428 GGTGATCGTCCCAATCGTGTACTACTCTGGTATTGGAAGCAACAGAGCTGATAGG 487
QY 140 MetIleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTySer 159
DB 488 ATGATCCGAAACGAGAAATTTTCGCTCTATCGGCTCAAGAAACCCCTAGTTTCTATTCT 547
QY 160 GlyLysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyArgLeuProHis 179
DB 548 GGGAAAGCTCTCTTAAGGCAATAGAACCACTGGATTATGAAGATATCGTTGCGCA 607
QY 180 HisGluThrGluLysTyGlnLysAlaGluLeuSerLeuCysArgValTyLysArgPro 199
DB 608 CATGAACCTGAACGATATCAAAAGGCTGAGATATCGTTTGGCGTGTTGACAAAGAGCT 667
QY 200 GlyValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsnHisAsn 219
DB 668 GGAGTGGAGGACCATCTCTTCCTCGCTGCTGCTCCCAACGAGG----- 712
QY 220 SerSerThrSerSerArgLeuAlaLeuArgGlnGlnHisSerSerSerSerSerAsn 239
DB 713 -----CCATCTCTCGTCAAGGGCT 730
QY 240 HisSerAspAsnLeuAsnLeuAsnAsn-----AsnAsnIleAsn 252
DB 731 TCTCATGACATATATTAAGAGCACATGAAATGGTTTCAGCACACATCAAC 781
```

RESULT 8

US-10-437-963-88434
Sequence 88434, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping.
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 88434
LENGTH: 1433
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_87286C.1
US-10-437-963-88434

Alignment Scores:
Pred. No.: 3,43e-81 Length: 1433
Score: 812.00 Matches: 190
Percent Similarity: 57.5% Conservative: 49
Best Local Similarity: 45.7% Mismatches: 97
Query Match: 39.9% Indels: 81
DB: 7 Gaps: 11

US-10-780-703-2 (1-379) x US-10-437-963-88434 (1-1433)

```
QY 42 GluAlaAspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThrGlu 61
DB 97 GAGGTAGATGGGCATGAGGATGACCTGTGTATGCCCGGTTTCAGGTTCCACCCGAGGAG 156
QY 62 GluGluLeuLeuGluPheTyLeuArgArgLysValGluGlyLysArgPheAsnValGlu 81
DB 157 GAGGAGCTGATCGAGTTCTACTCGCGGGAAGGTGGAGGCAAGCGCTTCAACGTCGAG 216
QY 82 LeuIleThrPheLeuAspLeuTyArgTyArgProTropGluLeuPro----- 97
DB 217 CTCATCACCTTCTCGACCTCTACCGCTACGACCCCTCGGAGCTCCCGGAATTGAGCAT 276
QY 98 -----AlaMetAlaAlaIleGlyGluLysGluTrpTy 108
DB 277 CTAATTAACAGCTCGATCGGGGTGCAGCAATGGCGCGATAGGGAGAGAGGTGTTTC 336
QY 109 PheTyValProArgAspArgLysTyArgAsnGlyAspArgProAsnArgValThrThr 128
DB 337 TTCTAGTGCCTCGGACAGGAAGTACAGGAACCGGGACCGGCCGAAACCGGTGACGGG 396
QY 129 SerGlyTyTrpLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgPro 148
DB 397 TCGGGGTACTGGAAGCGACCGGGGCGCATCGGATGATCCGAGCGAGAACCAACCGCCCC 456
QY 149 IleGlyLeuLysValPheTySerGlyLysAlaProLysGlyThrArgThr 168
DB 457 ATCGGGCTCAAGAACCGCTTGTCTTCTACTCCGCAAGGCCCCCAAGGGCTCGCAGC 516
QY 169 SerTrpIleMetAsnGluTyArgLeuProHisHisGluThrGluLysTyGlnLysAla 188
DB 517 AGCTGATCATGACGAGTATCGCTTCCCCCGCGCACCGACCGCTTACCACAGACT 576
QY 189 GluIleSerLeuCysArgValTyLysArgProGlyValGluAsp---HisProSerVal 207
DB 577 GAAATCTCACTATCGCGCTGTACAAAGCGCACCGGCATCGACGACGGCCATGGCCAGGTA 636
QY 208 -ProArg-----SerLeuSerThrArgHisAsnHisAsnSe 220
DB 637 TCCACGCGCGGTCTCGCGCACCTCCCGCGCGCGCGCGCACCGGTACAGGCAAC 696
QY 220 rSerThrSerArgLeuAlaLeuArgGln-----GlnGlnHisHisSerSe 237
DB 697 AAGCAAGGCTCTCTCGTCAGCTCCACGCGCGCGCTCCCAACTCCGCTCCCAAGCTCCAC 756
QY 237 rSerAsnHisSerAspAsnAsnLeuAsnAsnAsnIleAsnAsnLeuGluLysLe 257
DB 757 CTCCTCAGCAGCAGTGCACGTCGCGCGCGCGCATCTGCACGCGACCGCCCATGGTG 816
QY 257 userThrGluTySerGlyAspGlySerThrThrThrThrThrThrThrThrThrThr 277
DB 817 GCGCAACAGGACCGCTCGCGCGCG-----ACCATCAGCAGCAGCAGCAGCTCCACGCT 870
QY 277 pValThrIleAlaLeuAlaAsnGlnAsnIleTyArg-----PrometProTyAspTh 295
```


; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 81769
 ; LENGTH: 743
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(743)
 ; OTHER INFORMATION: unsure at all n locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_81262C.1
 ; US-10-437-963-81769

Alignment Scores:
 Pred. No.: 4,23e-78 Length: 743
 Score: 781.00 Matches: 139
 Percent Similarity: 89.5% Conservative: 15
 Best Local Similarity: 80.8% Mismatches: 18
 Query Match: 38.4% Indels: 0
 DB: 7 Gaps: 0

US-10-780-703-2 (1-379) x US-10-437-963-81769 (1-743)

| | | | |
|----|-----|--|-----|
| QY | 44 | AspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThrGluGlu | 63 |
| DB | 178 | GACGGCAGCGCAGCAGATGGTGATGCCGGGTTCCGGTTCACCCGACGGAGGAG | 237 |
| QY | 64 | LeuileGluPheThrLeuArgLysValGluGlyLysArgPheAsnValGluLeuile | 83 |
| DB | 238 | CTCATCGAGTTCTACCTCCGGCGAAGGTGGAGGGAAGCGTTCAACATCGAGCTCATC | 297 |
| QY | 84 | ThrPheLeuAspLeuThrArgTyArgProThrGluLeuProAlaMetAlaileGly | 103 |
| DB | 298 | GCCTTCGTGACCTCTACCGCTATGACCGGTGGGATCTCCCGCTCTGGCTCAATTGGG | 357 |
| QY | 104 | GluLysGluThrPheThrValProArgAspArgLysTyArgAsnGlyAspArgPro | 123 |
| DB | 358 | GACNAGAGTGGTCTCTATGTGCCNAGGACCGCATACCGAAGCGCGATCGGCC | 417 |
| QY | 124 | AsnArgValThrThrSerGlyTyThrLysAlaThrGlyAlaAspArgMetileArgSer | 143 |
| DB | 418 | AACCGAGTGACGCATCGGGGTACTGGAAGGCCACAGGGCGGATAGGATGGTGAAGTG | 477 |
| QY | 144 | GluThrSerArgProLleGlyLeuLysLysThrLeuValPheThrSerGlyLysAlaPro | 163 |
| DB | 478 | GAGGTGACCGCTCTATCGGCCTCAAGAACACGCTCTCTTACGTCCGCAAGGCGCCC | 537 |
| QY | 164 | LysGlyThrArgThrSerTrpilleMetAsnGluTyArgLeuProHisGluThrGlu | 183 |
| DB | 538 | AAGGGCTTCGACAGCTGGATCATGACAGATATGCTCTCTCTATGCGCAGCTGAC | 597 |
| QY | 184 | LysTyArgLysAlaGluileSerLeuLysArgValTyArgProGlyValGluAsp | 203 |
| DB | 598 | CGCTACCAAAAGGTAGAAATTCGTGTGTCGAGTCTACAAACGCCAGGAATCGAAGAC | 657 |
| QY | 204 | HisProSerValProArgSerLeuSerThrArgHis | 215 |
| DB | 658 | AACTGTAGTCGATCGAATCTCGATCTACCCGCTCAC | 693 |

RESULT 11
 US-10-425-115-173486
 ; Sequence 173486, Application US/10425115
 ; Publication No. US20040214272A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369326
 ; SEQ ID NO 173486
 ; LENGTH: 792
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: MRT4577_897C.1
 ; US-10-425-115-173486

Alignment Scores:
 Pred. No.: 4,52e-73 Length: 792
 Score: 737.00 Matches: 139
 Percent Similarity: 80.0% Conservative: 13
 Best Local Similarity: 73.2% Mismatches: 22
 Query Match: 36.2% Indels: 16
 DB: 8 Gaps: 2

US-10-780-703-2 (1-379) x US-10-425-115-173486 (1-792)

| | | | |
|----|-----|--|-----|
| QY | 41 | AspGluAlaAspAspHisAspMetValMetProGlyPheArgPheHisProThr | 60 |
| DB | 250 | GAGGCGGGGACGCGCAGCAACACGCTGGTGATGCCGGGTTCCGGTTCACCCACG | 309 |
| QY | 61 | GluGluGluLeuileGluPheThrLeuArgLysValGluGlyLysArgPheAsnVal | 80 |
| DB | 310 | GAGGAGGAGTGTCTACGTCTACCTCCGGCGCAAGTTGAGGGAAACGCTTCAACGTC | 369 |
| QY | 81 | GluLeuileThrPheLeuAspLeuThrArgTyArgProThrGluLeuProAlaMetAla | 100 |
| DB | 370 | GAGCTCATTTGCTTCTCGACCTCTACCGCTTCGACCCATGGGAGCTTCCAGCAATGGCG | 429 |
| QY | 101 | AlaileGlyGluLysGluThrPheThrValProArgAspArgLysTyArgAsnGly | 120 |
| DB | 430 | GTGATGGGCGGAGAGAGTGGTCTCTACGTCCGAGGACCGTAACTACCGAACGGA | 489 |
| QY | 121 | AspArgProAsnArgValThrThrSerGlyTyThrLysAlaThrGlyAlaAspArgMet | 140 |
| DB | 490 | GACCCGCCAAACCGGCTGACGGGTACTGGAAGCGACGGCGCTGACCGGATG | 549 |
| QY | 141 | IleArgSerGluThrSerArgProLleGlyLeuLysLysThrLeuValPheThrSerGly | 160 |
| DB | 550 | ATCCGAGGCGAGAACAAACCGCCCTATCGGGCTGGAAGAACACGCTCGTGTCTACTCCGC | 609 |
| QY | 161 | LysAlaProLysGlyThrArgThrSerTrpilleMetAsnGluTyArgLeuProHisHis | 180 |
| DB | 610 | AAGGCGCCCAAGGGCGTCCGACAGCTGGATCATGACAGTACCGCTTCCGCCACCA | 659 |
| QY | 181 | GluThrGluLys-----TyrgLysAlaGluileSerLeuLysArgValTyLys | 197 |
| DB | 670 | CCACCGAGCGCCGCCATTTGATTCCTCCAGTCCGAGATCTCTCTCGCGGCTCTACAAG | 729 |
| QY | 198 | ArgProGlyValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsn | 217 |
| DB | 730 | CGCTCAGGCATCGACGAC-----GGC | 750 |
| QY | 218 | HisAsnSerSerThrSerArgLeuAla | 227 |
| DB | 751 | CAGGCGAGTCTATCTCTCAGCACCCCAAGCT | 780 |

RESULT 12
 US-10-437-963-53771/c
 ; Sequence 53771, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.

113 QY gAspArgLysTyrArgAsnGlyAspArgProAenArgValThrThrSerGlyTyrTrp 133
 469 Db AGAAGAGAGTACCCAAAGTGGCGAGGCCAAACAGCGCTGCACACTTACGGGTATTGGAA 528
 133 QY sAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLys 153
 529 Db AGCCACGGGTACAGATAAGCCATATAGTAGTGAACCTCAAAAG---GTTGGGGTGAAGAA 585
 153 QY sThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMetAs 173
 586 Db ATCTTTGGTTTCTATGGAGGAAGCCACCAAAAGGGGTCAAAACTGATTCGATCATGCA 645
 173 QY nGluTyrArgLeuProHis-----Hi 180
 646 Db TGAGTACCGGTGTACTGAAACCAAGCCTAACATAGCCCTCTGGGTGTGACTTGGGCCA 705
 180 QY sGluThrGlyLysTyrGlnLysAlaGluIleSerLeuCysArgValTyrLysArgPro 200
 706 Db CAAGAAAACCTCCCTAAGGCTGGATGATTTGGGTGTGTCGGGATCTACAAGAGGCCAA 765
 200 QY yValGlu---AspHisPro-----SerValProAr 209
 766 Db CACACAAAGGTACATGAGAGGGATGATTCATGGATGATGATGGAGAAGTACTCTCC 825
 209 QY gSer-----LeuSerThrArgHisHisAsnHisAsnSerSerThrSerSe 224
 826 Db TTCCATCAACGTGGGCCACATAGTCGAGTGTTCACCTTCAAGATGTCACAGAGCTA 885
 224 QY rArgLeuAlaLeuArgGlnGlnHisHisSerSerSer-----SerAsnHi 240
 886 Db CAGTGGTGCATTTTGGAAATACACAGAAACACGTTAGAGGGGTGTATAGGCAATGG 945
 240 QY sSerAspAsnAsnLeuAsnAsnAsnLeuAsnLeuAsnLeuLysLeuSerThrG 260
 946 Db TAGTGTGAATGGATCAGCAATACTAATAATATTACAAGTGTCTATCATCTCATCA 1005
 260 QY uTyrSerGlyAspGlySerThrThrThrThrThrThrThrThrThrThrThrThr 280
 1006 Db GTTTGGC-----ACCTCAAACTCCAGGCAGAGCTT-----1036
 280 QY eAlaLeuAlaAsnGlnAsnIleTyrArgProMetProTyrAspThrSerAsnAsnThrLe 300
 1037 Db -----CTTCCTTTTGTCCCTTCAACACACACTTC 1065
 300 QY uIleValSerThrArg-----AsnHi 307
 1066 Db TAATTGGCTCCCAAGAAACACTCTCATCACTCTATTGGAATGTTGATGATGATTAATGA 1125
 307 QY sGlnAspAsp-----AspGluThrAlaIleValAspLeuGlnAr 321
 1126 Db TGATGATGATACAAACACTTCATTTGGACAGATAATGGAAATGTGAGTGTGTGGAGAC 1185
 321 QY gLeuValAsnTyrGlnIleSerAspGly-----AlaThrThrLeu-----334
 1186 Db TGNAGAGATAATAATAATACTAATGGTACTTCTGGCTCTTTGGCCACTCTGCTTAACCA 1245
 335 QY -MetProGlnThrGlnAlaLeuAlaMetAsnMetIleProAlaGlyThrIleProAs 354
 1246 Db GCTTCACAAACACCTTCATTCGCCCAATAATAGATGG-GTTACTTCCGAACCATATCAAA 1304
 354 QY nAsnAlaLeuTrpAspMetTrp 361
 1305 Db TACAAGGACCAATGGTATGG 1326

RESULT 14

US-10-425-114-20520
 ; Sequence 20520, Application US/10425114
 ; Publication No. US2004003488A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21153313/B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 20520
 ; LENGTH: 1644
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: LIB3242-006-H7_FLI
 US-10-425-114-20520

Alignment Scores:
 Pred. No.: 1,06e-49 Length: 1644
 Score: 535.00 Matches: 129
 Percent Similarity: 50.9% Conservative: 73
 Best Local Similarity: 32.5% Mismatches: 130
 Query Match: 26.3% Indels: 66
 DB: 7 Gaps: 12

US-10-780-703-2 (1-379) x US-10-425-114-20520 (1-1644)

QY 14 SerAsnGlnValAsnAsnAsnGluLysGlyIleGluAspAsnAspHis-ArgGlyGly 33
 Db 90 AACACATAGTAGAACACCAAC-----CTTGTAATAATAACCTGAGAGACCCCA 140
 QY 33 nGluSerHisValGlnAsnGluAspGluAlaAspAspHisAspHisAspMetValMetPr 53
 Db 141 AGACTCATCCAGGATTCACAGCAACACCCGAC-----CTGCCACC 182
 QY 53 oGlyPheArgPheHisProThrGluGluGluLeuIleGluPheTyrLeuArgArgLysVa 73
 Db 183 CGGGTTTCGGTTCACCCCAACAGAGAGCTCGTGGTTTCATTACCTCAAGAGAAAGT 242
 QY 73 lGluGlyLysArgPheAsnValGluLeuIleThrPheLeuAspLeuTyrArgTyrAspPr 93
 Db 243 AGATTCTGTCCTCCCTTCATCATCTGATGTTGATCTCTACAGTTTGTATCC 302
 QY 93 oTrpGluLeuProAlaMetAlaIleGlyLysGluTyrTrpTyrPheTyrValProAr 113
 Db 303 ATGGGAATTTACAGCAACAGCATCTGTTGGGCGGAGAGTGTACTTTTCAGCCCGAG 362
 QY 113 gAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrpLy 133
 Db 363 AGAAAGGAAGTACCAGATGGCGGAGGCCAAACAGGGCTGCACACTTCAGGGTATTGGAA 422
 QY 133 sAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLysLy 153
 Db 423 AGCCACGGGACAGATAAGCCATATAGTAGTGAACCTCAAAAG---GTTGGGGTGAAGAA 479
 QY 153 sThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMetAs 173
 Db 480 ATCTTTGGTTTCTATGGAGGAAGCCACCAAAAGGGGTAAAACCGATTCGATCATGCA 539
 QY 173 nGluTyrArgLeuProHis-----Hi 180
 Db 540 TGAGTACCGGTGTACTGAAACCAAGCCTAACATAGCCCTCTGGGTGTGACTTGGGCCA 599
 QY 180 sGluThrGlyLysTyrGlnLysAlaGluIleSerLeuCysArgValTyrLysArgProGl 200
 Db 600 CAAGAAAAAATCCCTTAAGGCTGATGATGGGTGTGTGTCGCCGAATCTACAAGAGAGGCA 659
 QY 200 yValGlu---AspHisPro-----SerValProAr 209
 Db 660 CACACAAAGGTACATGAGAGGAGCATTCATTGATGATGATGATGGAGAGTACTCTCC 719
 QY 209 gSer-----LeuSerThrArgHisHisAsnHisAsnSerSerThrSerSe 224

Db 720 TTCCATCAACGCGGCCACATGAGTGCAGAGTTTCCACCTTTCAAGGATGTCACGAGCTA 779
 Qy 224 rArgLeuAlaLeuArgGlnGlnGlnHisHisSerSerSer-----SerAsnHi 240
 Db 780 CAGTGGTCATTTGGGAAATGACAGAAACACGTTAGAGGGGTGTTATAGGCAATGG 839
 Qy 240 sSerAspAsnAsnLeuAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsn 260
 Db 840 TAGTGTGATGGATCAGCAATCACTAATAATATTAACAAGTGTATCATCATCTCATCA 899
 Qy 260 uTyrSerGlyAspGlySer-----ThrThrThrThrThrThrThrThrThrThr 272
 Db 900 GTTTGGCACCTCAAACTCAAGCGAGAGCTTCTCTTTGTCCTTCAAAACACATTC 959
 Qy 272 rAsnSerAsnSerAspValThrThrThrThrThrThrThrThrThrThrThrThrThr 290
 Db 960 TAATTCGGCCCTCCAAAGAACACTCTCATCTATTGGGATGTTGATGATCAAAACA 1019
 Qy 290 oMetProTyrAspThrSerAsnAsnThrLeuLeuValSerThrArgAsnHisGlnAspAs 310
 Db 1020 CTTCAATTTGACAGTAAATGGGATGCTGAGTGTGTTGAGGACAGATCAAGAGAATAATA 1079
 Qy 310 pAspGluThrAlaLeuValAspAspLeuGlnArgLeuValAsnTyrGlnIleSerAspG 330
 Db 1080 TAATAATAATGATCTTCTGGCTCCTTTGCTACTCTGCTTAACACG----- 1125
 Qy 330 yAlaThrThrLeuMetProGlnThrGlnAlaLeuAlaLeuAlaMetAsnMetIleProAlaG 350
 Db 1126 -----CTTCACAAACACCTTCAATGCCCCAAATAGATGG-GTTACTTCGGA 1171
 Qy 350 yThrIleProAsnAsnAlaLeuTipAspMetTipAsnProIleValPro 366
 Db 1172 CACCGTATCAATACACAGGACCAATGATGG-----TTAATTCGA 1214

RESULT 15

US-10-225-066A-849
 ; Sequence 849, Application US/10225066A
 ; Publication No. US20030226173A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mendel Biotechnology, Inc.
 ; APPLICANT: RATCLIFFE, Oliver
 ; APPLICANT: RIECHMANN, Jose Luis
 ; APPLICANT: ADAM, Luc J
 ; APPLICANT: DUBELL, Arnold T
 ; APPLICANT: HEARD, Jacqueline E
 ; APPLICANT: PILGRIM, Marsha L
 ; APPLICANT: JIANG, Cai-Zhong
 ; APPLICANT: REUBER, T. Lynne
 ; APPLICANT: CREELMAN, Robert A
 ; APPLICANT: PINEDA, Omaira
 ; APPLICANT: YU, Guo-Liang
 ; APPLICANT: BROUN, Pierre E
 ; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
 ; FILE REFERENCE: MB10036-2 US
 ; CURRENT FILING DATE: 2002-08-09
 ; PRIOR APPLICATION NUMBER: US/10/225,066A
 ; PRIOR FILING DATE: 2002-08-09
 ; PRIOR APPLICATION NUMBER: 09/837,444
 ; PRIOR FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: 60/310,847
 ; PRIOR FILING DATE: 2001-08-09
 ; PRIOR APPLICATION NUMBER: 60/336,049
 ; PRIOR FILING DATE: 2001-12-05
 ; PRIOR APPLICATION NUMBER: 60/338,692
 ; PRIOR FILING DATE: 2001-12-11
 ; PRIOR APPLICATION NUMBER: 10/171,468
 ; PRIOR FILING DATE: 2002-06-14
 ; NUMBER OF SEQ ID NOS: 1122
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 849
 ; LENGTH: 1130
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-10-225-066A-849

Alignment Scores:
 Pred. No.: 4,29e-49 Length: 1130
 Score: 527.50 Matches: 112
 Percent Similarity: 54.5% Conservative: 50
 Best Local Similarity: 37.7% Mismatches: 90
 Query Match: 25.9% Indels: 45
 DB: 6 Gaps: 8
 US-10-780-703-2 (1-379) x US-10-225-066A-849 (1-1130)
 Qy 37 ValGlnAsnGluAspGluAlaAspAspHisHis-----AspMetValMetProGly 54
 Db 33 GTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 92
 Qy 55 PheArgPheHisProThrGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 74
 Db 93 TTTAGGTTCCATCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 152
 Qy 75 GlyLysArgPheAsnValGluLeuLeuThrPheLeuAspLeuTyrArgTyrAspProTyr 94
 Db 153 'CACAAATCTCTTCCCTTTGATCTCATCAAGAAAGTCGACATTTACAATACGATCCATGG 212
 Qy 95 GluLeuProAlaMetAlaAlaIleGlyGlyLysGluTyrPheTyrPheTyrValProArgAsp 114
 Db 213 GACCTCCCAAGCTTTCAGCGATGGGGGAAAAAGAGTGTACTTTTATTTGCTTAGAGAC 272
 Qy 115 ArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTyrLysAla 134
 Db 273 AGGAAATACGCGAACACAGCAACAGCACTAACCGAGTAATCGAGGTGGCTTCTGAAAGCA 332
 Qy 135 ThrGlyAlaAspArgMetIleArgSer---GluThrSerArgProIleGlyLeuLysLys 153
 Db 333 ACCGAAACAGACCGGCTTATATCTGATGCTTCTGATGCTTCTGATGCTTCTGATGCTT 392
 Qy 154 ThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTyrPheMetAsn 173
 Db 393 TCACCTGTGTCTACCGTGTGAGCTGCTAAAGAGTCAAAACCGATGATGATGATGATGAT 452
 Qy 174 GluTyrArgLeuPro-----HisHisGluThr----- 182
 Db 453 GAATTCGTCTCCCTTCTCTCTGATGCTTCTGATGCTTCTGATGCTTCTGATGCTTCTGAT 512
 Qy 183 -----GluLysTyrGlnLysAlaGluIle 190
 Db 513 AAGAGCAACACCTTACAATAACAACAACAGCAGGAGCTTCTTCAACGATGCTTGG 572
 Qy 191 SerLeuCysArgValTyrLysArgPro---GlyValGluAspHisProSerValProArg 209
 Db 573 GCGATATGATATATTAAAGACAAATGCAGTATCTCTCACAAAGATCAATCCACAA 632
 Qy 210 SerLeu-----SerThrArgHisHisAsnHisAsnSerSerThrSer 223
 Db 633 TCTTGGGTTTATCCACGATTCCTGACAAATCAACAGTCAACACACACACACACGCAACT 692
 Qy 224 SerArgLeuAlaLeuArgGlnGlnGlnHisHisSerSerSerSerSerSerSerSer 243
 Db 693 CTCTTAGCTTCATCAGACGTTCTCAGCCATATATCAACAAGACAAACTTTATTCCTTCT 752
 Qy 244 AsnLeuAsnAsnAsnAsnAsnAsnAsn-----LeuGluLys 256
 Db 753 CCAGTCAACGACCGCGAGCTTCAAGATCAGTCTTCTTACTTTCGCTCTCAGATG 812
 Qy 257 LeuSerThrGluTyr-----SerGlyAspGlySerThrThrThrThrThrThrThr 271
 Db 813 CTGGAGTCAGTACATATACAGCAGAAACACGAAACAGGAGGATGCTCTGTTTCTGAGA 872
 Qy 272 ThrAsnSerAsnSerAspValThrIleAlaLeuAlaAsnGlnAsnIleTyr 288
 Db 873 AACAAATGGAACAGGGGATGCTCTGGTTCTGAGCAACAAATGAGATAACTAC 923

Search completed: April 6, 2006, 12:19:19

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2006, 09:13:07 ; Search time 1510 Seconds
(without alignments)

11502.112 Million cell updates/sec

Title: US-10-780-703-3

Perfect score: 2606

Sequence: 1 atggcaattgtatctccac.....ctaattcttttaagtaa 2606

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq 21:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*
- 14: Geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 2606 | 100.0 | 2606 | 14 | ADW23849 |
| 2 | 2606 | 100.0 | 2606 | 14 | Adx17043 Long Vege |
| 3 | 687.2 | 26.4 | 1444 | 10 | ADD30083 Plant yie |
| 4 | 687.2 | 26.4 | 1444 | 10 | Ades1520 Plant yie |
| 5 | 687.2 | 26.4 | 1444 | 12 | Adi44012 Plant tra |
| 6 | 467.2 | 17.9 | 1140 | 14 | ADW23847 |
| 7 | 467.2 | 17.9 | 1140 | 14 | ADW23847 Arabidops |
| 8 | 164 | 6.3 | 1212 | 8 | Adx17041 Long Vege |
| 9 | 164 | 6.3 | 1212 | 11 | ADA70597 Rice gene |
| 10 | 149 | 5.7 | 1250 | 8 | AC126812 Rice abio |
| 11 | 132.6 | 5.1 | 1549 | 8 | ADA71117 Rice gene |
| 12 | 131.8 | 5.1 | 1563 | 14 | ADA70311 Rice gene |
| 13 | 120.2 | 4.6 | 824 | 12 | Adw18167 Pinus rad |
| 14 | 115 | 4.4 | 879 | 12 | Ado62519 Transcrip |
| 15 | 114.2 | 4.4 | 1297 | 13 | Ado62518 Transcrip |
| 16 | 112.2 | 4.3 | 945 | 6 | Adx4580 Plant ful |
| 17 | 112.2 | 4.3 | 945 | 8 | ABZ13738 Arabidops |
| 18 | 112 | 4.3 | 1680 | 3 | ADA68463 Arabidops |
| 19 | 111.6 | 4.3 | 1635 | 13 | AAC36983 Arabidops |
| | | | | | Adx15275 Plant ful |

| | | | | | | |
|----|-------|-----|------|----|----------|--------------------|
| 20 | 111.6 | 4.3 | 1688 | 12 | ADO62516 | Ado62516 Transcrip |
| 21 | 110.4 | 4.2 | 1319 | 10 | ADD30271 | Add30271 Plant yie |
| 22 | 110.4 | 4.2 | 1319 | 12 | AD143646 | Adi43646 Plant tra |
| 23 | 110.4 | 4.2 | 1319 | 14 | AEA27112 | Aea27112 Stress to |
| 24 | 110.4 | 4.2 | 1606 | 3 | AAC42175 | Aac42175 Arabidops |
| 25 | 110.4 | 4.2 | 1638 | 3 | AAC44458 | Aac44458 Arabidops |
| 26 | 110.4 | 4.2 | 1641 | 3 | AAC45486 | Aac45486 Arabidops |
| 27 | 110.2 | 4.2 | 1359 | 13 | ADX45355 | Adx45355 Plant ful |
| 28 | 110.2 | 4.2 | 1580 | 13 | ADX45919 | Adx45919 Plant ful |
| 29 | 109.6 | 4.2 | 1384 | 12 | ADO62520 | Ado62520 Transcrip |
| 30 | 109.6 | 4.2 | 1398 | 13 | ADX54955 | Adx54955 Plant ful |
| 31 | 109.6 | 4.2 | 1557 | 12 | ADO62522 | Ado62522 Transcrip |
| 32 | 108.8 | 4.2 | 1445 | 13 | ADX31859 | Adx31859 Plant ful |
| 33 | 108.6 | 4.2 | 894 | 9 | ADB23141 | ADB23141 Environme |
| 34 | 108.6 | 4.2 | 894 | 12 | ADI44144 | Adi44144 Plant tra |
| 35 | 108.6 | 4.2 | 894 | 12 | ADO02466 | Ado02466 Thalecres |
| 36 | 108.6 | 4.2 | 894 | 12 | ADO61748 | Ado61748 Transcrip |
| 37 | 108.6 | 4.2 | 894 | 14 | AEA26816 | Aea26816 Stress to |
| 38 | 108.6 | 4.2 | 903 | 3 | AAC43036 | Aac43036 Arabidops |
| 39 | 108.6 | 4.2 | 923 | 12 | ADO63190 | Ado63190 Transcrip |
| 40 | 108.6 | 4.2 | 1125 | 10 | ADP30774 | Add30774 Plant yie |
| 41 | 108.6 | 4.2 | 1125 | 12 | ADI41634 | Adi41634 Plant tra |
| 42 | 108.6 | 4.2 | 1125 | 14 | AEA27084 | Aea27084 Stress to |
| 43 | 108.6 | 4.2 | 1234 | 3 | AAC50693 | Aac50693 Arabidops |
| 44 | 108.6 | 4.2 | 1237 | 3 | AAC33463 | Aac33463 Arabidops |
| 45 | 108.6 | 4.2 | 1260 | 10 | ADD55715 | Add55715 Thalecres |

ALIGNMENTS

| | | | | | | |
|----------|---|--|--|--|--|--|
| RESULT 1 | | | | | | |
| ADW23849 | | | | | | |
| ID | ADW23849 standard; DNA; 2606 BP. | | | | | |
| XX | AC | | | | | |
| AC | ADW23849; | | | | | |
| XX | 10-MAR-2005 (first entry) | | | | | |
| DT | XX | | | | | |
| XX | Arabidopsis thaliana LOV1 genomic sequence, SEQ ID 3. | | | | | |
| DE | XX | | | | | |
| XX | Flowering; plant; gene; ds. | | | | | |
| KW | XX | | | | | |
| XX | Arabidopsis thaliana. | | | | | |
| OS | XX | | | | | |
| XX | XX | | | | | |
| FH | Key | | | | | |
| FT | CDS | | | | | |
| FT | Location/Qualifiers | | | | | |
| FT | 1..2606 | | | | | |
| FT | /tag= b | | | | | |
| FT | /product= "LOV1 protein" | | | | | |
| FT | exon | | | | | |
| FT | 1..292 | | | | | |
| FT | /tag= a | | | | | |
| FT | /number= 1 | | | | | |
| FT | intron | | | | | |
| FT | 293..622 | | | | | |
| FT | /tag= c | | | | | |
| FT | /number= 1 | | | | | |
| FT | exon | | | | | |
| FT | 623..892 | | | | | |
| FT | /tag= d | | | | | |
| FT | /number= 2 | | | | | |
| FT | intron | | | | | |
| FT | 893..1923 | | | | | |
| FT | /tag= e | | | | | |
| FT | /number= 2 | | | | | |
| FT | /cons_splice= (5'site:NO,3'site:NO) | | | | | |
| FT | exon | | | | | |
| FT | 1924..2351 | | | | | |
| FT | /tag= f | | | | | |
| FT | /number= 3 | | | | | |
| FT | intron | | | | | |
| FT | 2352..2456 | | | | | |
| FT | /tag= g | | | | | |
| FT | /number= 3 | | | | | |
| FT | /cons_splice= (5'site:NO,3'site:NO) | | | | | |
| FT | exon | | | | | |
| FT | 2457..2606 | | | | | |
| FT | /tag= h | | | | | |
| FT | /number= 4 | | | | | |
| XX | XX | | | | | |

| | | | |
|----|------|---|------|
| Qy | 1681 | ATAAAAATAAATAATAATCAATCTTAAATTATATTTGCAATGTTACGTCAATGATAGT | 1740 |
| Db | 1681 | ATAAAAAATAAATAATAATCAATCTTAAATTATATTTGCAATGTTACGTCAATGATAGT | 1740 |
| Qy | 1741 | ACTCTAAATTTCTTCTAAACGCTGATCTTTTGGTAAATGCTTAACTTTACATAGTTTG | 1800 |
| Db | 1741 | ACTCTAAATTTCTTCTAAACGCTGATCTTTTGGTAAATGCTTAACTTTACATAGTTTG | 1800 |
| Qy | 1801 | TGAATCTTCTTTTCAAAACCATATCTTCGATAAAATGAATTTTTTCATAGATATTTGTAAGTC | 1860 |
| Db | 1801 | TGAATCTTCTTTTCAAAACCATATCTTCGATAAAATGAATTTTTTCATAGATATTTGTAAGTC | 1860 |
| Qy | 1861 | TATATTTGATAAATTTTGATATATGATATCAAGTCTCTTAATCAATGTGCTCATGTATAAATTAT | 1920 |
| Db | 1861 | TATATTTGATAAATTTTGATATATGATATCAAGTCTCTTAATCAATGTGCTCATGTATAAATTAT | 1920 |
| Qy | 1921 | AGGCTGAAATATCATCTGCGCGAGTGTAACAAAGCGCAGGAGTAGAAGATCATCCATCGG | 1980 |
| Db | 1921 | AGGCTGAAATATCATCTGCGCGAGTGTAACAAAGCGCAGGAGTAGAAGATCATCCATCGG | 1980 |
| Qy | 1981 | TACCACGTTCTCTCTCCACAAGACATCATTAACCAATAACTCATCGACATCATCCGTTTAG | 2040 |
| Db | 1981 | TACCACGTTCTCTCTCCACAAGACATCATTAACCAATAACTCATCGACATCATCCGTTTAG | 2040 |
| Qy | 2041 | CTTTAAGACAACAAACACACCATTCATCTCTCTTAATCAATTCGCGACAAACAACTTAAACA | 2100 |
| Db | 2041 | CTTTAAGACAACAAACACACCATTCATCTCTCTTAATCAATTCGCGACAAACAACTTAAACA | 2100 |
| Qy | 2101 | ACAAACAAACATCAACAAATCTCGAAGGCTCTCCACCGAATATTCGCGCAGCGCAGCA | 2160 |
| Db | 2101 | ACAAACAAACATCAACAAATCTCGAAGGCTCTCCACCGAATATTCGCGCAGCGCAGCA | 2160 |
| Qy | 2161 | CAACAAACGACACAAACAAACGTAACCTCTGACGTTACATTCGCTCTAGCCAAATCAAAACA | 2220 |
| Db | 2161 | CAACAAACGACACAAACAAACGTAACCTCTGACGTTACATTCGCTCTAGCCAAATCAAAACA | 2220 |
| Qy | 2221 | TATATGTCGCAATGSCCTTTACGACACAAAGCAACACATGATGATGTCCTACGAGAAATC | 2280 |
| Db | 2221 | TATATGTCGCAATGSCCTTTACGACACAAAGCAACACATGATGATGTCCTACGAGAAATC | 2280 |
| Qy | 2281 | ATCAAGACGATGATGAACCTGCCATTTGTGACGATCTTTCAAAGACTAGTTAACTACCAA | 2340 |
| Db | 2281 | ATCAAGACGATGATGAACCTGCCATTTGTGACGATCTTTCAAAGACTAGTTAACTACCAA | 2340 |
| Qy | 2341 | TATCAGATGGAGGTAACATCAATCACCAGATCTTTTCAAATGCTTCAACGTTTCATCATTA | 2400 |
| Db | 2341 | TATCAGATGGAGGTAACATCAATCACCAGATCTTTTCAAATGCTTCAACGTTTCATCATTA | 2400 |
| Qy | 2401 | CTCAACAAACAAATGCTTAAACGAAACCATTTACAAATTTGGTGGCTGCGGCGACTACAGCGA | 2460 |
| Db | 2401 | CTCAACAAACAAATGCTTAAACGAAACCATTTACAAATTTGGTGGCTGCGGCGACTACAGCGA | 2460 |
| Qy | 2461 | CAACGCTAAATGCGCTCAAACTCAAGCGCGGTTAGCTATGAACATGATTCCTGCAAGGAACGA | 2520 |
| Db | 2461 | CAACGCTAAATGCGCTCAAACTCAAGCGCGGTTAGCTATGAACATGATTCCTGCAAGGAACGA | 2520 |
| Qy | 2521 | TTCCAAAACAAATGCTTTGTGGGATATGTGGAATCCAAATAGTACCAAGATGGAAAAACAGATC | 2580 |
| Db | 2521 | TTCCAAAACAAATGCTTTGTGGGATATGTGGAATCCAAATAGTACCAAGATGGAAAAACAGATC | 2580 |
| Qy | 2581 | ACTATACTAATATTTCCCTTTTAAAGTAA | 2606 |
| Db | 2581 | ACTATACTAATATTTCCCTTTTAAAGTAA | 2606 |

RESULT 2

ADX17043
ID ADX17043 standard: DNA: 2606 BP.

AC ADX17043;

XX
DT 21-APR-2005 (first entry)

| | |
|----|--|
| XX | Long Vegetative phase 1, LOV1, gene, SEQ ID 3. |
| DE | Long Vegetative phase 1; LOV1; flowering; plant; gene; ds; |
| XX | Arabidopsis thaliana. |
| OS | |
| XX | |
| FH | Key Location/Qualifiers |
| FT | CDS 1..2606 |
| FT | /tag= b |
| FT | /product= "LOV1" |
| FT | 1..292 |
| FT | /tag= a |
| FT | /number= 1 |
| FT | intron 293..622 |
| FT | /tag= c |
| FT | /number= 1 |
| FT | exon 623..891 |
| FT | /tag= d |
| FT | /number= 2 |
| FT | intron 892..1922 |
| FT | /tag= e |
| FT | /number= 2 |
| FT | exon 1923..2352 |
| FT | /tag= f |
| FT | /number= 3 |
| FT | intron 2353..2457 |
| FT | /tag= g |
| FT | /number= 3 |
| FT | exon 2458..2606 |
| FT | /tag= h |
| FT | /number= 4 |
| XX | |
| PN | US2005034194-A1. |
| XX | |
| PD | 10-FEB-2005. |
| XX | |
| PF | 19-FEB-2004; 2004US-00780703. |
| XX | |
| PR | 20-FEB-2003; 2003KR-00010772. |
| XX | |
| PA | (UYSE-) UNIV SEOUL NAT IND FOUND. |
| XX | |
| PI | Lee JS, Kim YH, Choi EK, Yoo SY, Ahn JH, Choi YD; |
| XX | |
| DR | WPI; 2005-054880/15. |
| DR | P-ESDB; ADX17042. |
| XX | |
| PT | New LOV1 polynucleotide encoding a polypeptide controlling |
| PT | time of plants, i.e., either delaying or inducing early fl |
| PT | plants. |
| XX | |
| PS | Claim 2; SEQ ID NO 3; 28pp; English. |
| XX | |
| CC | The present invention relates to Arabidopsis thaliana Long |
| CC | phase 1, LOV1, gene (ADX17043), coding sequence (ADX17041) |
| CC | (ADX17042). LOV1 controls the flowering time of plants by |
| CC | activity of flowering-promoting gene AGL20, and so is used |
| CC | controlling flowering: the flowering time of plants can be |
| CC | LOV1 gene overexpression, or early flowering can be induc |
| CC | suppressing the expression of the LOV1 gene. The LOV1 gene |
| CC | from Arabidopsis thaliana by an activation tagging screen. |
| XX | |
| SQ | Sequence 2606 BP; 896 A; 509 C; 388 G; 813 T; 0 U; 0 Other |

Query Match
Best Local

Matches 260

45

Db 1

PA (JIAN/) JIANG C.
 PA (HEAR/) HEARD J. E.
 PA (HAK/) HAKE V.
 PA (CREE/) CREELMAN R. A.
 PA (RATC/) RATCLIFFE O.
 PA (ADAM/) ADAM L. J.
 PA (REUB/) REUBER T. L.
 PA (KEDD/) KEDDIE J.
 PA (BROU/) BROUN P. E.
 PA (PILG/) PILGRIM M. L.
 PA (DUBE/) DUBELL A. N.
 PA (PINE/) PINEDA O.
 PA (YUGG/) YU G.
 XX
 PI Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;
 PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Broun PE;
 PI Pilgrim ML, Dubell AN, Pineda O, Yu G;
 XX
 DR WPI; 2004-132245/13.
 DR P-PSDB; ADI44013.
 XX
 PT New transgenic plant comprising a recombinant polynucleotide of any one
 PT of more than 500 nucleotide sequences, useful in bioinformatic search
 PT methods.
 XX
 PS Disclosure; SEQ ID NO 2475; 435pp; English.
 XX
 CC The invention describes a transgenic plant comprising a recombinant
 CC polynucleotide of any one of more than 500 nucleotide sequences fully
 CC defined in the specification or its complement. The method of the
 CC invention can be used to produced a plant having altered traits such as:
 CC enhanced tolerance to abiotic stress; glycosylase tolerance; hormone
 CC sensitivity; disease resistance; sugar sensing; early or late flowering;
 CC altered flower structure, change in stem bifurcations, altered branching
 CC pattern, reduced apical dominance, reduced trichome density; lack of
 CC trichomes; reduced ectopic trichome development; altered trichome
 CC development; increase in trichome number; altered stem morphology;
 CC increased root growth; increased root hairs; altered seed development;
 CC premature senescence; increased necrosis; increase in seedling or plant
 CC size; decreased plant size; leaf morphology; seed morphology; seed
 CC biochemistry; increase in root anthocyanins; increase in plant
 CC anthocyanins, or alteration in light response or shade avoidance. The
 CC transgenic plant, polynucleotides and polypeptides are useful in
 CC bioinformatic search methods. This sequence represents a plant
 CC transcription factor related polynucleotide.
 XX
 SQ Sequence 1444 BP; 490 A; 350 C; 246 G; 358 T; 0 U; 0 Other;
 Query Match 26.4%; Score 687.2; DB 12; Length 1444;
 Best Local Similarity 96.8%; Pred. No. 26-127;
 Matches 701; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
 Qy 1883 GTATCAAGTCTCTAATCAATGTGCTCATGTATATATTAGGCTGAATATCATTTGTGCGG 1942
 Db 669 GTATCGTCTTCGGACCATGAAACCGAGAGTAGTACCAAGAGCTGAATATCATTTGTGCGG 728
 Qy 1943 AGTGTAAGAAAGGCGAGAGTAGAAGATCATTCATCGGTACACGTTCTCTCTCCACAAG 2002
 Db 729 AGTGTAAGAAAGGCGAGAGTAGAAGATCATTCATCGGTACACGTTCTCTCTCCACAAG 788
 Qy 2003 ACATCATACCAATCAATCATTCGACATCATCCGTTTAGCCTTAGACACACACCA 2062
 Db 789 ACATCATACCAATCAATCATTCGACATCATCCGTTTAGCCTTAGACACACACCA 848
 Qy 2063 TTCACTCTCTCTAATCATTCGACACCAACCTTTAAACAAACAAACATCAACATCT 2122
 Db 849 TTCACTCTCTCTAATCATTCGACACCAACCTTTAAACAAACAAACATCAACATCT 908
 Qy 2123 CGAGAAGCTCTCCACCGCAATATTCGGGCGAGGCGAGCAACAAACAAACCAACAG 2182
 Db 909 CGAGAAGCTCTCCACCGCAATATTCGGGCGAGGCGAGCAACAAACAAACCAACAG 968

Qy 2183 TAACTCTGACGTATACATTGCTCTAGCCAAATCAAAACATATATCGTCCAAATGCCCTTACGA 2242
 Db 969 TAACTCTGACGTATACATTGCTCTAGCCAAATCAAAACATATATCGTCCAAATGCCCTTACGA 1028
 Qy 2243 CACAAGCAACAAACATTTGATGTCTCTA CGAGAAATCATCAAGACGATGATGAAACTGC 2302
 Db 1029 CACAAGCAACAAACATTTGATGTCTCTA CGAGAAATCATCAAGACGATGATGAAACTGC 1088
 Qy 2303 CATTGTTGACGATCTTCAAGACATAGTTAACTACCAATATACAGATGGAGGTAAACATCAA 2362
 Db 1089 CATTGTTGACGATCTTCAAGACATAGTTAACTACCAATATACAGATGGAGGTAAACATCAA 1148
 Qy 2363 TCACCAATACTTTCAAAATTTGCTCAACAGTTTTCATCATCTCAACAAACAAATGCTAAAGC 2422
 Db 1149 TCACCAATACTTTCAAAATTTGCTCAACAGTTTTCATCATCTCAACAAACAAATGCTAAAGC 1208
 Qy 2423 AAACGATTAACAATTTGGTGGCTGGGCGACTACAGCGACAAAGCTTAATGCTCAAAACTCA 2482
 Db 1209 AAACGATTAACAATTTGGTGGCTGGGCGACTACAGCGACAAAGCTTAATGCTCAAAACTCA 1268
 Qy 2483 AGCGGCGTTAGCTATGACATGATTCCTGCAAGGACGATTCCTCAACAAACATGCTTTGTGGGA 2542
 Db 1269 AGCGGCGTTAGCTATGACATGATTCCTGCAAGGACGATTCCTCAACAAACATGCTTTGTGGGA 1328
 Qy 2543 TATGTGGAATCCAATAGTACCAGATGGAAACAGAGATCACTATATATTCCTTTTAA 2602
 Db 1329 TATGTGGAATCCAATAGTACCAGATGGAAACAGAGATCACTATATATTCCTTTTAA 1388
 Qy 2603 GTAA 2606
 Db 1389 GTAA 1392
 RESULT 6
 ADW23847
 ID ADW23847 standard; cDNA; 1140 BP.
 XX
 AC ADW23847;
 XX
 DT 10-MAR-2005 (first entry)
 XX
 DE Arabidopsis thaliana LOV1 coding sequence, SEQ ID 1.
 XX
 KW Flowering; plant; gene; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 Key Location/Qualifiers
 CDS 1..1140
 FT /*tag= a
 FT /product= "LOV1 protein"
 XX
 FN KR2004075252-A.
 XX
 PD 27-AUG-2004.
 XX
 PF 20-FEB-2003; 2003KR-00010772.
 XX
 PR 20-FEB-2003; 2003KR-00010772.
 XX
 PA (UYSE-) UNIV SEOUL NAT.
 XX
 PI Ahn JH, Choi EG, Choi YD, Kim YH, Lee JS, Yoo SY;
 XX
 DR WPI; 2005-054880/06.
 XX
 DR P-PSDB; ADW23848.
 XX
 PT New LOV1 polynucleotide encoding a polypeptide controlling the flowering
 PT time of plants, i.e., either delaying or inducing early flowering of the
 PT plants.
 XX
 PS Claim 4; SEQ ID NO 1; 25pp; Korean.
 XX

CC The present invention relates to the novel gene LOV1 (Long Vegetative phase 1) from Arabidopsis thaliana and its use in controlling flowering time in plants. LOV1 inhibits the flowering stimulating gene AGL20. LOV1 can be used for manipulating flowering time by over-expressing LOV1 to delay flowering, or inhibiting expression of LOV1 to stimulate flowering. CC LOV1 is also useful for identifying homologous genes in other plants. The CC present sequence is the LOV1 coding sequence.

XX Sequence 1140 BP; 393 A; 272 C; 216 G; 259 T; 0 U; 0 Other;

Query Match 17.9%; Score 467.2; DB 14; Length 1140;
Best Local Similarity 82.3%; Pred. No. 1.7e-83;
Matches 596; Conservative 0; Mismatches 23; Indels 105; Gaps 1;

QY 1883 GTATCAAGTCTTAATCAATGCTCTATGTATATTAATTAAGCTGAATATCATTTGGCGG 1942
DB 522 GTATCGTCTTCCGACCATGAAACCGAGAGTAGTACCAAAAGGCTGAATATCATTTGGCGG 581
QY 1943 AGTGTAACAAAGCCAGGAGTAGAAGATCATCCATCGGTACCAAGTCTCTCTCCACAAG 2002
DB 582 AGTGTAACAAAGCCAGGAGTAGAAGATCATCCATCGGTACCAAGTCTCTCTCCACAAG 641
QY 2003 ACATCAATACCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 2062
DB 642 ACATCAATACCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 701
QY 2063 TTATCT 2122
DB 702 TTATCT 761
QY 2123 CGAAGAAGTCTTCCACCGAATATTCCGCGGAGCGGAGCAGCAGCAACCAACCAACAG 2182
DB 762 CGAAGAAGTCTTCCACCGAATATTCCGCGGAGCGGAGCAGCAGCAACCAACCAACAG 821
QY 2183 TAACTCTGAGTTCATGCTCTAGCCAAATCAAAATCATATATCGTCCAAATCTTACGA 2242
DB 822 TAACTCTGAGTTCATGCTCTAGCCAAATCAAAATCATATATCGTCCAAATCTTACGA 881
QY 2243 CACAAACCAACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2302
DB 882 CACAAACCAACACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 941
QY 2303 CATTTGTTGAGGATCTTCAAGAGCTAGTTAACTACCAAAATCATAGATGAGGTAACATCA 2362
DB 942 CATTTGTTGAGGATCTTCAAGAGCTAGTTAACTACCAAAATCATAGATGAGGTAACATCA 989
QY 2363 TCACCAATATCTTTCAAAATGCTCAACAGTTTTCATATCATCTCAACAAACAAATGCTAACGC 2422
DB 990 ----- 989
QY 2423 AAACGCAATTACAATTGGTGGCTGGGGGACTACAGGAGCAACCGCTTAATGCTCAAACTCA 2482
DB 990 -----AGCGACAAACCGCTTAATGCTCAAACTCA 1016
QY 2483 AGCGGGTTAGCTATGAACATGATTTCTCGAGGAACGATTTCCAAACAAATGCTTTGGGGA 2542
DB 1017 AGCGGGTTAGCTATGAACATGATTTCTCGAGGAACGATTTCCAAACAAATGCTTTGGGGA 1076
QY 2543 TATGTGGAATCCCAATAGTACAGATGGAACAGAGATCACTATCAATATTCCTTTTAA 2602
DB 1077 TATGTGGAATCCCAATAGTACAGATGGAACAGAGATCACTATCAATATTCCTTTTAA 1136
QY 2603 GTAA 2606
DB 1137 GTAA 1140

RESULT 7
ADX17041
ID ADX17041 standard; cDNA; 1140 BP.
XX
AC
ADX17041;
XX

DT 21-APR-2005 (first entry)
XX Long Vegetative phase 1, LOV1, coding sequence, SEQ ID 1.
XX Long Vegetative phase 1, LOV1; flowering; plant; gene; ss.
XX Arabidopsis thaliana.
XX Key Location/Qualifiers
FT CDS 1..1140
FT /*tag= a
FT /product= "LOV1"
XX PN US2005034194-A1.
XX PD 10-FEB-2005.
XX PF 19-FEB-2004; 2004US-00780703.
XX PR 20-FEB-2003; 2003KR-00010772.
XX PA (UYSE-) UNIV SEOUL NAT IND FOUND.
XX Lee JS, Kim YH, Choi EK, Yoo SY, Ahn JH, Choi YD;
PI WPI: 2005-054880/15.
DR P-PSDB; ADX17042.
XX New LOV1 polynucleotide encoding a polypeptide controlling the flowering
PT time of plants, i.e., either delaying or inducing early flowering of the
PT plants.
XX Claim 2; SEQ ID NO 1; 28pp; English.
XX The present invention relates to Arabidopsis thaliana Long Vegetative
CC phase 1, LOV1, gene (ADX17043), coding sequence (ADX17041) and protein
CC (ADX17042). LOV1 controls the flowering time of plants by repressing the
CC activity of flowering-promoting gene AGL20, and so is useful for
CC controlling flowering: the flowering time of plants can be delayed by
CC LOV1 gene overexpression, or early flowering can be induced by
CC suppressing the expression of the LOV1 gene. The LOV1 gene was isolated
CC from Arabidopsis thaliana by an activation tagging screen.
XX Sequence 1140 BP; 393 A; 272 C; 216 G; 259 T; 0 U; 0 Other;

Query Match 17.9%; Score 467.2; DB 14; Length 1140;
Best Local Similarity 82.3%; Pred. No. 1.7e-83;
Matches 596; Conservative 0; Mismatches 23; Indels 105; Gaps 1;
QY 1883 GTATCAAGTCTTCTAATCAATGCTCTATGTATATTAATTAAGCTGAATATCATTTGGCGG 1942
DB 522 GTATCGTCTTCCGACCATGAAACCGAGAGTAGTACCAAAAGGCTGAATATCATTTGGCGG 581
QY 1943 AGTGTAACAAAGCCAGGAGTAGAAGATCATCCATCGGTACCAAGTCTCTCTCCACAAG 2002
DB 582 AGTGTAACAAAGCCAGGAGTAGAAGATCATCCATCGGTACCAAGTCTCTCTCCACAAG 641
QY 2003 ACATCAATACCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 2062
DB 642 ACATCAATACCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 701
QY 2063 TTATCT 2122
DB 702 TTATCT 761
QY 2123 CGAAGAAGTCTTCCACCGAATATTCCGCGGAGCGGAGCAGCAGCAACCAACCAACAG 2182
DB 762 CGAAGAAGTCTTCCACCGAATATTCCGCGGAGCGGAGCAGCAGCAACCAACCAACAG 821
QY 2183 TAACTCTGAGTTCATGCTCTAGCCAAATCAAAATCATATATCGTCCAAATCTTACGA 2242
DB 822 TAACTCTGAGTTCATGCTCTAGCCAAATCAAAATCATATATCGTCCAAATCTTACGA 881

2243 CACAAGCAACACACATTGATGCTCTACGAGAAATCATCAAGACGATGATGAAACTGC 2302
 882 CACAAGCAACACACATTGATGCTCTACGAGAAATCATCAAGACGATGATGAAACTGC 941
 2303 CATTGTTGACGATCTTCAAGACTGTTAACTACCAAAATATCAGATGGAGGTAACATCAA 2362
 942 CATTGTTGACGATCTTCAAGACTGTTAACTACCAAAATATCAGATGG----- 989
 2363 TCACCAATACTTTTCAAATGCTCAACAGTTTCATCATACTCAACAAATGCTTAACGC 2422
 990 ----- 989
 2423 AAACGCAATTACAATTGGTGGCTGCGCGGACTACAGCAACACGCTTAATGCTCAAACTCA 2482
 990 -----AGCGCAACGCTTAATGCTCAAACTCA 1016
 2483 ACCGCGGTTAGCTATGAACATGATTCCTGCGAGAACGATTCCAACAAATGCTTTGGGA 2542
 1017 ACCGCGGTTAGCTATGAACATGATTCCTGCGAGAACGATTCCAACAAATGCTTTGGGA 1076
 2543 TATGTGGAATCCAATAGTACAGATGGAAACAGAGATCACTATCTAATATTTCTTTTAA 2602
 1077 TATGTGGAATCCAATAGTACAGATGGAAACAGAGATCACTATCTAATATTTCTTTTAA 1136
 2603 GTAA 2606
 1137 GTAA 1140

RESULT 8

ADA70597
 ID ADA70597 standard; DNA; 1212 BP.
 XX
 AC ADA70597;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Rice gene, SEQ ID 3920.
 XX
 KW Plant; bacterial infection; fungal infection; viral infection; rice;
 KW gene; ds.
 XX
 OS Oryza sativa.
 XX
 PN WO200300898-A1.
 XX
 PD 03-JAN-2003.
 XX
 PF 22-JUN-2001; 2001WO-IB001105.
 XX
 PR 22-JUN-2001; 2001WO-IB001105.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
 XX
 DR WPI; 2003-175290/17.
 XX

Identifying at least one gene involved in plant resistance or response to
 pathogenic infection for conferring resistance or tolerance to a plant to
 bacterial, fungal or viral infection by determining or detecting plant
 gene expression.

Claim 6; SEQ ID NO 3920; 899pp; English.

The present invention relates to a method (M1) for identifying genes
 involved in plant resistance or response to pathogenic infection. M1
 comprises identifying a gene whose expression is significantly altered in
 the incompatible interaction of plant gene expression relative to
 expression of the gene in an uninfected plant, in a mutant plant that
 does not express a gene associated with response to pathogenic infection,
 or in a corresponding incompatible or compatible interaction. (M1) is

CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 XX

Sequence 1212 BP; 245 A; 421 C; 361 G; 185 T; 0 U; 0 Other;
 Query Match 6.3%; Score 164; DB 8; Length 1212;
 Best Local Similarity 70.8%; Pred. No. 6.2e-23;
 Matches 218; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 605 AGATATATGATATGACGCTATGCGCGGATAGAGAGAAAGAGTGTACTTCTATGTGC 664
 DB 161 AGCTCGATGCGGGGTGACGCAATGCGCGGATAGGGGAGAGAGTGTCTTCTACGTGC 220
 QY 665 CAAGAGATCGGAAATATAGAAATGAGATAGACCGAACCGAGTACGACTTCAGGATATT 724
 DB 221 CTCGGGACAGAGATACAGGAAACCGGACCGGCGGATGACGGGTGCGGGGTACT 280
 QY 725 GGAAAGCCACCGGAGCTGATAGATGATCAGATCGGAGACTTCTCGGCTATCGGATTA 784
 DB 281 GGAAGGCGACGGGGCGGACCGGATGATCCGAGCGGAGAACACCGCCCATCGGGCTCA 340
 QY 785 AGAAACCCCTAGTCTTCTACTCTGCTAAAGCCCTAAAGGCACTGCTACTAGTTGGATCA 844
 DB 341 AGAAGACGCTTGTCTTCTACTCTCGGCAAGGCCCCCAAGGGCGTCCGACGAGCTGGATCA 400
 QY 845 TGAAGAGTATCGTCTTCCGACCATGAAACCGGAGAGTACCAAAAGGTATAAATTCTAC 904
 DB 401 TGAAGAGTATCGTCTTCCGACCATGAAACCGGACCGGCTACCAAGACTGAAATCTCAC 460
 QY 905 TATAACTC 912
 DB 461 TATCCGC 468

RESULT 9

ACL26812
 ID ACL26812 standard; cDNA; 1212 BP.
 XX
 AC ACL26812;
 XX
 DT 02-JUN-2005 (first entry)
 XX
 DE Rice abiotic stress responsive polynucleotide SEQ ID NO:768.
 XX
 KW ss; abiotic stress tolerance; transgenic plant; plant; cereal;
 KW agriculture.
 XX
 OS Oryza sativa.
 XX
 PN WO2003008540-A2.
 XX
 PD 30-JAN-2003.
 XX
 PF 21-JUN-2002; 2002WO-US019668.
 XX
 PR 22-JUN-2001; 2001US-0300112P.
 PR 24-AUG-2001; 2001US-0314682P.
 PR 26-SEP-2001; 2001US-0325277P.
 PR 21-NOV-2001; 2001US-0332132P.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;
 PI Moughamer T, Provart N, Ricke D, Zhu T;
 XX
 DR WPI; 2003-248011/24.
 XX

New stress-responsive nucleic acid, useful for altering the
 responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
 stress, salt stress or osmotic stress.
 Claim 1; SEQ ID NO 768; 89pp; English.

XX The invention relates to novel abiotic stress responsive polynucleotides
 CC and polypeptides. Also disclosed are vectors, expression cassettes, host
 CC cells, and plants containing such polynucleotides. Also disclosed are
 CC methods for using the polynucleotides and polypeptides to alter the
 CC responsiveness of a plant to abiotic stress. The invention is useful in
 CC agriculture. The nucleic acid is useful for determining whether a test
 CC plant has been exposed to an abiotic stress condition. It is also useful
 CC for selecting an agent that alters abiotic stress regulated
 CC polynucleotide expression in a plant cell, and to identify a homolog or
 CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
 CC molecule and the polypeptide encoded by it are useful in altering the
 CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
 CC stress, osmotic stress or any of their combinations. The present sequence
 CC is used in the exemplification of the invention

XX SQ Sequence 1212 BP; 245 A; 421 C; 361 G; 185 T; 0 U; 0 Other;
 Query Match 6.3%; Score 164; DB 11; Length 1212;
 Best Local Similarity 70.8%; Pred. No. 6.2e-23;
 Matches 218; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 605 AGATATATGTATATGACGATATGCGCGCGATAGGAGAGAAAGTGGTACTTCTATGTGC 664
 DB 161 AGCTCGATGGGGTGCAGCAATGGCGCGATAGGGGAGAGAGTGGTCTTCTACGTGC 220
 QY 665 CAAGAGATCGGAATATAGAAATGGAGATAGACCGAACCGAGTAACGACTTCAGGATAT 724
 DB 221 CTCGGGACAGGAAGTACAGGAACCGGGCCGCGGAGCGGTGACGGGTCTCGGGTACT 280
 QY 725 GGAAGCCACCGGAGTGTATGATGATCAGATCGGAGACTTCTCGGCTATCGGATTA 784
 DB 281 GGAAGCGACGGGGCGGACCGGATGATCGAGCCGAGAACACCCGCCATCGGCTCA 340
 QY 785 AGAAACCCCTAGTGTCTTCTACTCTGTGTAAGCCCTTAAAGGCACTCGTACTAGTTGGATCA 844
 DB 341 AGAAGACGCTTGTTCTTCTACTCTCGGCAAGGCCGCCCAAGGGCGTCCGACGCTGGATCA 400
 QY 845 TGACGAGTATCGTCTTCTCGCACCATGAACCGAGAGTACCAAGGATTAATTTCTAC 904
 DB 401 TGAACGAGTATCGCTTCTCGCCCGGACACCGCGCTACCAAGAGCTGAAATCTCAC 460
 QY 905 TATACTC 912
 DB 461 TATGCCGC 468

RESULT 10
 ADA71117
 ID ADA71117 standard; DNA; 1250 BP.
 AC ADA71117;
 XX 20-NOV-2003 (first entry)
 DT Rice gene, SEQ ID 4440.
 XX Rice gene, SEQ ID 4440.
 DE Plant; bacterial infection; fungal infection; viral infection; rice;
 KW gene; ds.
 XX Oryza sativa.
 OS Oryza sativa.
 XX WO2003000898-A1.
 FN WO2003000898-A1.
 PD 03-JAN-2003.
 XX 22-JUN-2001; 2001WO-IB001105.
 XX 22-JUN-2001; 2001WO-IB001105.
 PR (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
 XX WPI; 2003-175290/17.
 DR Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 XX Claim 6; SEQ ID NO 4440; 899pp; English.
 XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

XX SQ Sequence 1250 BP; 291 A; 388 C; 339 G; 230 T; 0 U; 2 Other;
 Query Match 5.7%; Score 149; DB 8; Length 1250;
 Best Local Similarity 71.1%; Pred. No. 6.2e-20;
 Matches 197; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 618 TGCAGCTATGCGCGGATAGGAGAGAAAGTGGTACTTCTATGTGCCAAGAGTCGGAA 677
 DB 141 TCCCGCTCTGGCCTCAATTGGGGACAGAGTGGTCTTCTATGTGCCAAGAGTCGGAA 200
 QY 678 ATATAGATGGAGATAGACCGAACCGAGTACGACTTCAGGATATGGAAGCCACCGG 737
 DB 201 GTACCGAAACGGCGATCGGCCCAACCGAGTACGCGGTACTCGAAGCCACAGG 260
 QY 738 AGCTGATAGGATGATCAGATCGGAGACTTCTCGGCTATCGGATTAAGAAACCCCTAGT 797
 DB 261 GCGGATAGGATGGTGAAGTGGAGGTGACCGTCTATCGCCTCAAGAGACGCTCGT 320
 QY 798 TTTCTACTCTGTAAAGCCCTTAAAGGCACCTCGTACTAGTTGGATCATGAACGAGTATCG 857
 DB 321 CTTCTACGTCGCAAGGGCGCCCAAGGGCTTCGCAAGCAGCTGGATCATGAACGAGTATCG 380
 QY 858 TCTTCGCGACCATGAACCGGAGTACCAAGGTA 894
 DB 381 TCTCCCTCATGGCGACGCTGACCGCTACCAAGGATA 417

RESULT 11
 ADA70311
 ID ADA70311 standard; DNA; 549 BP.
 AC ADA70311;
 XX 20-NOV-2003 (first entry)
 DT Rice gene, SEQ ID 3634.
 XX Rice gene, SEQ ID 3634.
 DE Plant; bacterial infection; fungal infection; viral infection; rice;
 KW gene; ds.
 XX Oryza sativa.
 OS Oryza sativa.
 XX WO2003000898-A1.
 FN WO2003000898-A1.
 PD 03-JAN-2003.
 XX 22-JUN-2001; 2001WO-IB001105.
 XX 22-JUN-2001; 2001WO-IB001105.
 PR (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

```
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
DR WPI; 2003-175290/17.  
XX Identifying at least one gene involved in plant resistance or response to  
PT pathogenic infection for conferring resistance or tolerance to a plant to  
PT bacterial, fungal or viral infection by determining or detecting plant  
PT gene expression.  
XX Claim 6; SEQ ID NO 3634; 899pp; English.  
XX The present invention relates to a method (M1) for identifying genes  
CC involved in plant resistance or response to pathogenic infection. M1  
CC comprises identifying a gene whose expression is significantly altered in  
CC the incompatible interaction of plant gene expression relative to  
CC expression of the gene in an uninfected plant, in a mutant plant that  
CC does not express a gene associated with response to pathogenic infection,  
CC or in a corresponding incompatible or compatible interaction. (M1) is  
CC useful for conferring resistance to resistance or tolerance to a plant to  
CC bacterial, fungal or viral infection. The present sequence was used to  
CC illustrate the invention.  
XX  
SQ Sequence 549 BP; 103 A; 174 C; 194 G; 78 T; 0 U; 0 Other;  
Query Match 5.1%; Score 132.6; DB 8; Length 549;  
Best Local Similarity 69.5%; Pred. No. 9.6e-17; Mismatches 79; Indels 0; Gaps 0;  
Matches 180; Conservative 0;  
QY 620 CAGCTATGCGCGCATAGGAGAGAAAGTGGTACTTCTATGTGCCAAGAGATCGGAAT 679  
DB 260 CCGCAATGGCGGTATAGGGAGAGAGTGGTCTTCTAGTCCGCGGACCGCAAGT 319  
QY 680 ATAGAAATGGAGATAGACCGAACCGAGTAACGACTTCAGAGATATGGAAACCCACCGAG 739  
DB 320 ACCGMAACGGACCGCGCGAACCGGTGACGGGTGCGGGTACTGGGAAGCGACGGGG 379  
QY 740 CTGATAGATGATCAGATCGGAGACTTCTCGGCTATCGGATTAAGAAACCTAGTTT 799  
DB 380 CCGACCGGATGATCCGCGGCGAGAACACGCCGCCCATCGGCTCAAGAAAGACGTCGTCT 439  
QY 800 TCTACTCTGGTAAAGCCCTAAAGCACTCCGTACTAGTTGGATCATGAACGAGTATGTC 859  
DB 440 TCTACTCGGCAAGGCCCCCAAGGGCGTCCGACGACGCTGATCATGACGAGTACGCT 499  
QY 860 TTCGCGACCAATGAACCGA 878  
DB 500 TCCCGCGCGCGCGCAGA 518  
RESULT 12  
ADW18167  
ID ADW18167 standard; cDNA; 1563 BP.  
XX  
AC ADW18167;  
XX  
DT 24-MAR-2005 (first entry)  
XX  
XX Pinus radiata transcription factor cDNA NAC family Seq 1946.  
XX  
KW gene; ss; plant; transcription; gene regulation; gene expression;  
KW transgenic plant; drought resistance; disease resistance; salt tolerance;  
KW cold tolerance; freezing tolerance; flowering; flavor enhancer;  
KW flower color.  
OS Pinus radiata.  
XX  
FN WO2005001050-A2.  
XX  
XX 06-JAN-2005.  
PD  
XX 07-JUN-2004; 2004WO-US017965.  
PF
```

```
XX 06-JUN-2003; 2003US-0476189P.  
XX (ARBO-) ARBORGEN LLC.  
XX  
PI Bloksberg LN, Bryant C, Connett MB, Emerson SJ, Frost MJ;  
PI Forster RLS, Grigor M, Higgins C, Lasham A, Lund ST, Maguin A;  
PI Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;  
XX WPI; 2005-075542/08.  
DR P-PSDB; ADW18488.  
XX  
XX New polynucleotides isolated from plants encoding transcription factors,  
PT and polypeptides encoded by such polynucleotides, useful for regulating  
PT gene transcription and gene expression.  
XX  
PS Claim 3; SEQ ID NO 1946; 1265pp; English.  
XX  
CC This invention relates to novel isolated plant nucleic acid molecules, or  
CC variants thereof, that encode transcription factors. Specifically, it  
CC refers to transcription factor proteins that are capable of binding to  
CC DNA in order to regulate gene transcription and gene expression in a  
CC plants, in particular Eucalyptus grandis and Pinus radiata. The present  
CC invention describes DNA constructs containing DNA encoding a  
CC transcription factor that regulates the promoter, which is operably  
CC linked to the desired nucleic acid to be expressed. It further provides  
CC transgenic plants expressing a transcription factor that confers a trait  
CC change, enhanced cold/frost tolerance, enhanced color, health and  
CC nutritional characteristics, as well as improved taste, starch  
CC composition, flower longevity and germination, amongst others.  
CC Accordingly, such plants that are successfully transfected with a DNA  
CC construct can be characterized by a difference in flower color, petal or  
CC leaf shape and size, aroma or plant height. This polynucleotide is a  
CC plant transcription factor cDNA sequence of the invention.  
XX  
SQ Sequence 1563 BP; 501 A; 280 C; 369 G; 413 T; 0 U; 0 Other;  
Query Match 5.1%; Score 131.8; DB 14; Length 1563;  
Best Local Similarity 65.4%; Pred. No. 1.8e-16;  
Matches 193; Conservative 0; Mismatches 102; Indels 0; Gaps 0;  
QY 618 TGCAGCTATGCGCGCATAGGAGAGAAAGTGGTACTTCTATGTGCCAAGAGATCGGAA 677  
DB 333 TCCAGCTTTGGCTTCAATATGAGAGAAAGATGTTCTTTTACGTACATCGAGATAGAA 392  
QY 678 ATATAGAAATGGAGATAGACCGAACCGAGTAACGACTTCAGGATATTTGGAAAGCCACCGG 737  
DB 393 GTATAGGAATGGAGATCGACCTAACAGAGTTACAAAGTCTGGGTATTTGGAAGCTACAGG 452  
QY 738 AGCTGATAGGATGATCAGATCGGAGACTTCTCGGCTATCGGATTAAGAAACCCCTAGT 797  
DB 453 TGTGACCGGATGGTTCACAGTGAAGTCTCTGGTTGCATCGGCTTGAAGAAACGCTTGT 512  
QY 798 TTTCTACTCTGGTAAAGCCCTTAAAGCACTCGTACTAGTTGGATCATGAACGAGTATCG 857  
DB 513 GTTCTATACAGGAGGCGGCCCAAGGAAAGGAAAGCAAGCTGGATCATGACCGAATATAG 572  
QY 858 TCTTCCGCACCATGAAACCGAGAGTACCAAAAGGTATATAATTTCTACTATTAATTC 912  
DB 573 GTTGCCACAACCTCGAAACTCGCAGAAATTCAGAAAGACGAACTCTCCCTTTGCCCG 627  
RESULT 13  
ADOG2519  
ID ADOG2519 standard; DNA; 824 BP.  
XX  
AC ADOG2519;  
XX  
XX 15-JUL-2004 (first entry)  
DT  
XX Transcription factor G1412 orthologous sequence, SEQ ID 986.  
DE  
XX
```

KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;
 KW osmotic stress tolerance; cold tolerance; heat tolerance;
 KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
 KW glyphosate resistance; flowering; fertility; seed development; ds.
 XX Glycine max.
 XX WO2004031349-A2.
 XX 15-APR-2004.
 XX 18-SEP-2003; 2003WO-US030292.
 XX 18-SEP-2002; 2002US-0411837P.
 XX 17-DEC-2002; 2002US-0434166P.
 XX 24-APR-2003; 2003US-0465809P.
 XX (MEND-) MENDEL BIOTECHNOLOGY INC.
 XX Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
 XX Riechmann JL, Haake V, Dubell AN, Keddle JS, Sherman BK;
 XX WPI; 2004-330163/30.
 XX New recombinant polynucleotide encoding transcription factor
 PT polypeptides, useful for producing transgenic plants with advantageous
 PT properties compared to a reference plant.
 PS Claim 1; SEQ ID NO 986; 510pp; English.
 XX The present invention relates to novel plant transcription factor
 CC proteins (I) and nucleotide sequences (II) (AD061534-AD063778). The
 CC sequences can be used to produce transgenic plants, which overexpress
 CC (II), where the transgenic plant has an altered trait as compared to a
 CC non-transgenic plant or wild-type plant. The transgenic plant comprises
 CC an altered trait selected from increased tolerance to abiotic stress,
 CC increased tolerance to osmotic stress, increased tolerance to cold,
 CC increased germination in cold, increased tolerance to heat, increased
 CC germination in heat, increased tolerance to freezing conditions,
 CC increased tolerance to low nitrogen conditions, increased tolerance to
 CC low phosphate conditions, increased tolerance to disease, including
 CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
 CC increased tolerance to multiple fungal pathogens, increased resistance to
 CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
 CC increased sensitivity to ACC, altered sugar sensing, increased tolerance
 CC to sugars, altered carbon/nitrogen sensing, early flowering, late
 CC flowering, altered flower structure, loss of flower determinacy, reduced
 CC fertility, altered shoot meristem development, altered branching pattern,
 CC altered stem morphology, altered vascular tissue structure, reduced
 CC apical dominance, altered trichome density, altered trichome development,
 CC altered trichome structure, altered root development, altered shade
 CC avoidance, altered seed development, altered seed ripening, altered seed
 CC germination, slow growth, fast growth, altered cell differentiation,
 CC altered cell proliferation, altered cell expansion, altered phase change,
 CC altered senescence, abnormal embryo development, altered programmed cell
 CC death, lethality when overexpressed, altered necrosis patterns, increased
 CC plant size, increased biomass, large seedlings, dwarfed plants, dark
 CC green leaves, change in leaf shape, increased leaf size and mass, light
 CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
 CC altered seed coloration, altered seed size, altered seed shape, large
 CC seed, increased leaf wax, increased seed fatty acids, altered seed oil
 CC content, altered seed protein content, altered seedprenyl content,
 CC altered leaf prenyl lipid content, increased anthocyanin levels, and
 CC decreased anthocyanin levels. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 824 BP; 254 A; 143 C; 154 G; 272 T; 0 U; 1 Other;
 Query Match 4.6%; Score 120.2; DB 12; Length 824;
 Best Local Similarity 53.1%; Pred. No. 3.2e-14;
 Matches 372; Conservative 0; Mismatches 289; Indels 39; Gaps 4;

QY 157 CCGGATTTAGATTCCTCCAGGAGAGAACTCATAGAGTTTACCTTCGGCGAAA 216
 Db 164 CTTGATTTAGATTTTATCCCATGATGAGAGCTTTTGGTTAGTACCTTTGCGCGAAG 223
 QY 217 GTTGAAGGCAACGCTTTAATGTAGAACTCATCACTTTTCTCGATCTTTATCGTATGAT 276
 Db 224 GTTCTGGCCATCATTTCTCTTCCCAATCATCTGAGTTGATTTGTACAAGTTTGTAT 283
 QY 277 CTTTGGGAACCTTCTGTTAAATATATATCATTAACAACACACATAAATCATCTCAAACTA 336
 Db 284 CCATGGGTTCTTCCAGGT-AAATTTATTTTCATTTATCATATAAAAAAATCTAAAAACG 342
 QY 337 TTTGGAATCTTAATTTCTATTCATATGTTAAGATCTTTCTCTCTCTATCACCTTTCTC 396
 Db 343 ATCTTTAACCAAGATTTTGTCAATATCAAAAGTTTGTAGAGTTGCGAATCACAAATTCG 402
 QY 397 TCTCTATTTCTTTTTTAACTTATATGTACTACCTCTCTCTTATGAAGTATTACTATG 456
 Db 403 GACCGCTTAGATCACAATTTGCTTCAATTTCTTCCCATAT----- 442
 QY 457 TCGATCGTTTAAACAATCTCAATATCTTTAAAGCTTCTCCCTCTTTAGTTTCTTTCTTAA 516
 Db 443 -----TAAAGATTTGAATATTAATTTGTAGTTGTGATTTACAACGGTAAATTTAAAAAATTT 495
 QY 517 ATTAACCTTAATTAACAACCTTACATATATATATATAAGATATACAAATATGTGTATGTTTT 576
 Db 496 ATAACCTTGGTTAATG-----TCTATTTGCAATGCACAAATAATGAGGACTACTAGG 547
 QY 577 CATATATAGCTTATGTATGTTTAAATCATAGATATATATATGATGATGATGATGATGATG 636
 Db 548 ATTTATTTATTTATTTTAAATTTGATTTGTTTATGTTGTTGTTGTTGTTGTTGTTGTTG 607
 QY 637 GGAGAGAAAGTGGTACTTCTATGTGCGCAAGATCGGAATATATAGAAATGCGATAGA 696
 Db 608 GGAGAGAAAGTGGTACTTCTTCCAGTCCAAAGACAGGAAGTACCCGAATGGTTTCACGA 667
 QY 697 CCGAACCGAGTAACGACTTCAGGATATTGGAAGCCACCGAGCTGTATAGGATGATCAGA 756
 Db 668 CCAACAGAGTTGCGGGTTCTGGGTTATTTGGAAGCCACTGGAACTGAACTGAACTATCACC 727
 QY 757 TCGGAGACTTCTCGGCCCTTATCGGATTTAAAGAAAAACCCCTAGTTTCTTCTCTGTTAAAGCC 816
 Db 728 ACTGA---AGGTAGAAAAAGTTGGCATAAAAAAGCACTTGTTTCTACGTTGGCAAGCA 784
 QY 817 CTTAAAGCACTCTGACTAGTTGATCATGAAACGATATC 856
 Db 785 CCCAAAGGCTCCAAACCCATTGGATCATGCACGATATC 824
 RESULT 14
 AD062518
 ID AD062518 standard; DNA; 879 BP.
 XX
 AC AD062518;
 XX
 DT 15-JUL-2004 (first entry).
 XX
 DE Transcription factor G1412 orthologous sequence, SEQ ID 985.
 KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;
 KW osmotic stress tolerance; cold tolerance; heat tolerance;
 KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
 KW glyphosate resistance; flowering; fertility; seed development; ds.
 XX Glycine max.
 XX OS
 XX WO2004031349-A2.
 XX 15-APR-2004.
 XX 18-SEP-2003; 2003WO-US030292.

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2006, 09:13:28 ; Search time 9249 Seconds
(without alignments)
13182.737 Million cell updates/sec

Title: US-10-780-703-3

Perfect score: 2606

Sequence: 1 atggcaattgtatctccac.....ctaataattctttaagtaa 2606

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 segs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*
10: gb_est10.*
11: gb_est11.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 466 | 17.9 | 507 | 8 | AY201074 |
| 2 | 450 | 17.3 | 901 | 8 | DR749910 |
| 3 | 432.8 | 16.6 | 506 | 1 | AI993449 |
| 4 | 395.8 | 15.2 | 690 | 9 | BH481683 |
| 5 | 333 | 12.8 | 838 | 9 | BH666975 |
| 6 | 323.2 | 12.4 | 622 | 1 | AV827617 |
| 7 | 314.2 | 12.1 | 947 | 8 | DR749909 |
| 8 | 292.8 | 11.2 | 427 | 1 | AV440745 |
| 9 | 292.8 | 11.2 | 468 | 1 | AV442690 |
| 10 | 229 | 8.8 | 283 | 8 | R84139 |
| 11 | 226.2 | 8.7 | 729 | 9 | BZ028065 |
| 12 | 219.2 | 8.4 | 702 | 9 | BH999674 |
| 13 | 214.4 | 8.2 | 280 | 9 | AY201073 |
| 14 | 202.4 | 7.8 | 720 | 9 | BZ028439 |
| 15 | 185.4 | 7.1 | 869 | 10 | CZ981915 |
| 16 | 184.4 | 7.1 | 516 | 3 | BI893633 |
| 17 | 183.2 | 7.0 | 586 | 9 | CC955522 |
| 18 | 180 | 6.9 | 605 | 10 | AG223605 |
| 19 | 178.2 | 6.8 | 623 | 8 | CX527114 |
| 20 | 178.2 | 6.8 | 648 | 1 | AW683672 |
| 21 | 178.2 | 6.8 | 653 | 2 | BE204243 |
| 22 | 176.8 | 6.8 | 741 | 7 | CO096386 |

| | | | | | |
|----|-------|-----|------|----|----------|
| 23 | 172.4 | 6.6 | 813 | 6 | CD575840 |
| 24 | 171.4 | 6.6 | 525 | 8 | DT006901 |
| 25 | 171 | 6.6 | 718 | 10 | CM220628 |
| 26 | 169.8 | 6.5 | 174 | 11 | CR399335 |
| 27 | 167.6 | 6.4 | 813 | 9 | AV126200 |
| 28 | 166.4 | 6.4 | 310 | 1 | AV428075 |
| 29 | 164 | 6.3 | 1404 | 10 | CL960332 |
| 30 | 163.6 | 6.3 | 853 | 8 | DR928151 |
| 31 | 163.6 | 6.3 | 906 | 8 | DR935181 |
| 32 | 162 | 6.2 | 880 | 8 | DR938316 |
| 33 | 161 | 6.2 | 555 | 1 | AW443468 |
| 34 | 153.6 | 5.9 | 486 | 8 | DT004645 |
| 35 | 153.2 | 5.9 | 508 | 10 | CG269114 |
| 36 | 153.2 | 5.9 | 621 | 10 | CG136476 |
| 37 | 153.2 | 5.9 | 759 | 10 | CG192989 |
| 38 | 153.2 | 5.9 | 822 | 10 | CG178208 |
| 39 | 153.2 | 5.9 | 823 | 9 | CC659605 |
| 40 | 153.2 | 5.9 | 894 | 10 | CG178210 |
| 41 | 152.8 | 5.9 | 657 | 2 | BI268158 |
| 42 | 151 | 5.8 | 431 | 1 | AV798788 |
| 43 | 149.2 | 5.7 | 649 | 7 | CK077495 |
| 44 | 149.2 | 5.7 | 770 | 6 | CB621762 |
| 45 | 149.2 | 5.7 | 784 | 6 | CB621699 |

ALIGNMENTS

RESULT 1
AY201074

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AY201074 507 bp DNA linear GSS 03-JAN-2005
Arabidopsis thaliana Landsberg DNA Arabidopsis thaliana
genomic clone GT109.D5.08.02.99.b.507, genomic survey sequence.

AY201074

AY201074.1 GI:27897028

GSS.

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 507)

May B.P., Simorowski J., Arroyo, J.-M., Vaughn, M.W., Shen, R.,

McCombie, W.R. and Martienssen, R.A.

Arabidopsis genomic sequences flanking Ds enhancer and gene traps

in transgenic lines

Unpublished (2004)

Contact: Martienssen RA

1 Bungtown Rd., Cold Spring Harbor, NY 11724, USA

Tel: 516 367 8322

Fax: 516 367 8369

Email: martiens@cshl.org

Arabidopsis thaliana sequence flanking Ds end of Ds-Trap insertion

from line GT109.

Class: transposon-tagged.

Location/Qualifiers

1..507

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/cultivar="ecotype Landsberg"

/db_xref="taxon:3702"

/clone="GT109.D5.08.02.99.b.507"

/clone_lib="Arabidopsis thaliana Landsberg DNA"

ORIGIN

Query Match 17.9%; Score 466; DB 9; Length 507;

Best Local Similarity 98.4%; Pred. No. 1.3e-81;

Matches 492; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

QY 457 TCGATCGTTAACAAATTCCTCAATATCTTTAAACGCTTCTCCCTTTAGTTCTTTCTTAA 516

|||||

Db 10 TCGATCGTTAACAAATTCCTCAATATCTTTAAACGCTTCTCCCTTTAGTTCTTTCTTAA 69

|||||


```
COMMENT
Other GSSs: BOGNE16TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES
    source
    1..690
    /organism="Brassica oleracea"
    /mol_type="genomic DNA"
    /strain="TO1000DH3"
    /db_xref="taxon:3712"
    /clone="BOGNE16"
    /note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared
    genomic DNA inserted into pHOS1 using BstXI linkers"

ORIGIN
Query Match      15.2%; Score 395.8; DB 9; Length 690;
Best Local Similarity 80.3%; Pred. No. 1e-67;
Matches 561; Conservative 0; Mismatches 107; Indels 31; Gaps 7;

QY 1875 TGATATGATGATCAAGTCTCTTAATCATGCTGCTCATGTATA-ATTATAGGCTGAAATATC 1933
DB |||||
DB 19 TAATATATATGTCAGTCTCTGATCAATGTTGTTGATATATATATGCTGAAATATC 78
QY 1934 ATTTGGCGGAGTGTCACAAAGCCAGGAGTAGAAGATCATCCATCGGTACACGTTCTCT 1993
DB |||||
DB 79 ATTTGGCGGAGTGTCACAAAGCCAGGAGTAGAAGATCATCTATCTTACCAGGCTCTCT 138
QY 1994 CTCCAAGACATCATATAACATAATCATCTCGACATCATCCGTTTGGCTTAAAGACAACA 2053
DB |||||
DB 139 ATCCAAGACATCATATAATCATACCTCATCTCTCCGCTTGGCTTAAAGACAACAACA 198
QY 2054 ACAACACGATTCATCTCTTAATCATCTTCCAGACAACACCTTTAACAACAACAACAT 2113
DB |||||
DB 199 ACAACACCACTCATCATCAATCAATTCAGACAACAACCTTTAAACAACA----- 249
QY 2114 CAACAATCTCGAGAGCTCTCCACCGAATATTCCGCGGAGCGGACAGACAACAACGAC 2173
DB |||||
DB 250 CAACAATCTTGAGAGCTCTCAACCGAATATTTTGGCGAGCGGTAGC---AGCATACAC 306
QY 2174 CACAAACAGTAATCTGAGCTTACCATTTGCTTAGCCATCAAAACATATATCGTCCAT 2233
DB |||||
DB 307 CACAAACAGTATTCTGAGCTCACTATAGCTTTAGCCATCAAAACATTTATCGTCCAT 366
QY 2234 GCCTTAGCACAAGCAACAACATTTGATGTTCTTAGGAGAAATCATCAAGACGATGA 2293
DB |||||
DB 367 GCCTTAGGTGTAAGCAACACCAATGAT-----CTCCACAACCTTAATAAAGAGATGA 420
QY 2294 TGAACCTGCATTTGAGAGCTTCAAGAGCTAGTTAACTACCAATATCATAGATGGAG- 2352
DB |||||
DB 421 TGAACACGCAATGTTGATGATCTTCAAGAGCTCGTTAACTACCAATATCTGATGGAG 480
QY 2353 --GTAACATCAATCACCAATCTTTCAAAATTTGCTCAACAGTTTTCATCTACTCAACA 2410
DB |||||
DB 481 TAGTAACATCAATCACCAATCTATCAATCGCTCAACAGTTTTCATAT-----CAACA 534
QY 2411 AAATGCTAACGCAACGCAATTAATTTGTTGGTGGTGGCGGAGCTACAGCGCAACCGTAA 2470
DB |||||
DB 535 AGAGGTAAATCAACCGCTTGCATTTGGTGGCGGGGCAACTTACAGTG---GCGCTAAC 591
QY 2471 GCCTCAACCTCAAGCGGCTTACCTATGAACATGTTCTTGGAGGAGCAATTTCCAAACA 2530
DB |||||
DB 592 GCCTCAACCGCAGGGGCAATTAAGCATGATGATCCCTCGAGGAGCAATTTCCAAACAG 651
QY 2531 TGCCTTGTGGGATATGTGGAATCCAAATAGTACCAGATGG 2569
DB |||||
DB 652 TGCCTTGTGGGATCTATGGAATCAATCACTAGTACCAGATGG 690

RESULT 5
BH666975
LOCUS
DEFINITION
BOMKF40TR BO 2.3 KB Brassica oleracea genomic clone BOMKF40,
genomic survey sequence.
ACCESSION
BH666975
VERSION
BH666975.1 GI:18726179
KEYWORDS
GSS.
SOURCE
Brassica oleracea
ORGANISM
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
1 (bases 1 to 838)
Ayele,M., Haas,B.J., Wortman,J.R., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
Utterback,T.R., White,O.R. and Town,C.D.
Whole genome shotgun sequencing of Brassica oleracea and its
application to gene discovery and annotation in Arabidopsis
Genome Res. 15 (4), 487-495 (2005)
15805490
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
    1..838
    /organism="Brassica oleracea"
    /mol_type="genomic DNA"
    /strain="TO1000DH3"
    /db_xref="taxon:3712"
    /clone="BOMKF40"
    /clone_lib="BO 2.3 KB"
    /note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared
    genomic DNA inserted into pHOS1 using BstXI linkers"

ORIGIN
Query Match      12.8%; Score 333; DB 9; Length 838;
Best Local Similarity 80.1%; Pred. No. 2.7e-55;
Matches 507; Conservative 0; Mismatches 80; Indels 46; Gaps 8;

QY 1983 CCAGGTTCTCTCTCCACAGACATCAATACCATCACTCATCATCATCCGTTTAGCC 2042
DB |||||
DB 2 CCAGGTTCTACATCCACAGACATCAATACCATCACTCATCATCATCCGTTGGCC 61
QY 2043 TTAAAG-----ACAACAACAACACCATTCATCTCTCTTAATCATTCGACACAACCTT 2096
DB |||||
DB 62 GAATGACAAACAACAACAACACCATTCCTCTCTCTTAATCATTCGACACAACCTT 121
QY 2097 AACACAACAACAACATCAACAATCTCGAGAGCTCTCCACCGAATATTCCGGCGACGCG 2156
DB |||||
DB 122 AACAA-----CAACAATCTCGACAGCTCTCAACCGAATATTCCGGCGACGCT 169
QY 2157 AGCAACAACAACGACCAACAACAGTAACCTCTGACGTTTACCATTTGCTTAGCCAAATCAA 2216
DB |||||
DB 170 AGC---ACCATAACCACTACAAACAGTAATTTCTGACGTCACCATAGCTTTAGCCAAACAG 226
QY 2217 AACATATATGTCGAATGCCCTTAGACAACAACAACATTCATGATGCTCTTAGGAGA 2276
DB |||||
DB 227 AACATCTATGTCGAATGCCCTTTTGGTGCAGCAACACCAATTAAT-----A 274
QY 2277 AATCATCAAGACGATGATGAAGTGCCTTGTGACGATCTTCAAGACTAGTTAACTA- 2335
DB |||||
DB 275 TCTAATCAAGAGATGATGAAGACCGCAATAGTTGATGATCTTCAAGACTCGTTAACTAC 334
QY 2336 --CCAAATATCAGATGGAGGTAAACATCAATCAACCAATACTTTCAAATTTGCTCAACAGTTT 2393
DB |||||
```

335 GGCCTAATATCTGGTGGAGTAACATCAATCAACCAATCACTCAAAATGGCTCAACAGTTT 394
 2394 CATCATCTACACACAAATGCTTAACGCAACGCAATACATTTGGTGGCTGGCGGACT 2453
 395 CATATCAACAACTACTA-----AACGCAATGCGTTGCAATCGGTGACGCGCAGACT 448
 2454 ACAGCGCAACGCTAATGCTCAAACTCAAGCGGCTTACGTATGATGAAATGATTCCTGCA 2513
 449 ACAG--CAGGTTAACGCTCAACAGCGGCGACGTTAGCGATGAACATGATCCTGCA 505
 2514 GGAAGCATTCAAACAATGCTTTGGGATATGTGAATCCAAATAGTACCAGATGAAC 2573
 506 GGGAGCATTCAAACAATGCTTTGGGAGATGTGAATCCAAATGATGATGATGAGG-AAC 564
 2574 AGAGATCACTATCACTAATATTCCTTTTAAGTAA 2606
 565 AAGATCATTTACTAATATTCCTTTTAGATAA 597

RESULT 6
 AV827617
 LOCUS
 DEFINITION AV827617 RAF19 Arabidopsis thaliana cDNA clone RAFL09-17-121 5', mRNA sequence.
 mRNA sequence.

ACCESSION AV827617
 VERSION AV827617.1 GI:19869677
 KEYWORDS EST
 SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses.

REFERENCE 1 (bases 1 to 622)
 Seki, M., Nakazaki, K., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
 Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
 Arabawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
 and Shinozaki, K.

TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
 JOURNAL Unpublished (2002)
 COMMENT Contact: Motoaki Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060

Email: mseki@rtc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
 and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
 al., submitted for publication) digested with BamHI and SalI. This
 clone is in a modified pBluescript vector. Please visit our web
 site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
 details.

FEATURES
 source
 1..622
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAFL09-17-121"
 /dev stages="plants at various developmental stages from
 germination to mature seeds"
 /lab host="DH10B"
 /clone_lib="RAFL9"
 /note="Site 1: BamHI; Site 2: SalI; subjected to
 dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
 hr) treatments"

ORIGIN
 Query Match 12.4%; Score 323.2; DB 1; Length 622;
 Best local Similarity 93.6%; Pred. No. 2.3e-53;
 Matches 337; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

1883 GTATCAAGTCTCTAATCAATGCTCATGTATATATAGGCTGAATATCATTTGTCGCG 1942

263 GTATCGTCTTCGACCATGAACCGAGAGTAGTACCAAGGCTGAATATCATTTGTCGCG 322
 1943 AGTGTAACAAAGGCGAGAGTAGAAGATCATCCATCGGTACACAGTTCTCTCTCCACAAG 2002
 323 AGTGTAACAAAGGCGAGAGTAGAAGATCATCCATCGGTACACAGTTCTCTCTCCACAAG 382
 2003 ACATCATAACCAATCAATCATCGACATCATCCGTTTAGCCTTAAGACAAACACACACCA 2062
 383 ACATCATAACCAATCAATCATCGACATCATCCGTTTAGCCTTAAGACAAACACACACCA 442
 2063 TTCAATCTCTCTTAATCAATCCGACAAACCTTTAAACAAACACACACCAATCT 2122
 443 TTCAATCTCTCTTAATCAATCCGACAAACCTTTAAACAAACACACCAATCT 502
 2123 CGAGAAGCTCTCCACCGAATATTCGGCGAGCGGAGACACACACACACCAACACAG 2182
 503 CGAGAAGCTCTCCACCGAATATTCGGCGAGCGGAGACACACACACCAACACAG 562
 2183 TAACCTGACGTTACCATTTGCTTAGCCAAATCAAAACATATATGTCCTCAATGCTTAGCA 2242
 563 TAACCTGACGTTACCATTTGCTTAGCCAAATCAAAACATATATGTCCTCAATGCTTAGCA 622

RESULT 7
 DR749909
 LOCUS
 DEFINITION 85-L020254-065-002-E11-Sela MP1Z-ADIS-065d Arabidopsis thaliana
 cDNA clone 002-E11, mRNA sequence.

ACCESSION DR749909
 VERSION DR749909.1 GI:71035249
 KEYWORDS EST
 SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses.

REFERENCE 1 (bases 1 to 947)
 Paz-Ares, J., Valencia, A., Costantino, P., Vittorioso, P., Davies, B.,
 Gilmartin, P., Giraudat, J., Parcy, F., Reindl, A., Sablowski, R.,
 Coupland, G., Martin, C., Angenent, G.C., Baeumlein, H., Mock, H.P.,
 Carbonero, P., Colombo, L., Tonelli, C., Engstrom, P.,
 Droegge-Laser, W., Gatz, C., Kavanagh, T., Kushnir, S., Zabeau, M.,
 Laux, T., Hordsworth, M., Ruberti, I., Ratcliff, F., Smeekens, S.,
 Somssich, I., Weisshaar, B. and Traas, J.

RGII, an EU project on functional genomics of transcription
 factors from Arabidopsis thaliana
 Comp. Funct. Genomics 3 (2), 102-108 (2002)
 Contact:

Paz-Ares, Costantino, Vittorioso, Davies, Gilmartin, Giraudat, Parcy, Sablowski, Coupland, Martin, Angenent, Baeumlein, Carbonero, Colombo, Tonelli, Engstrom, Droegge-Laser, Gatz, Kavanagh, Kushnir, Zabeau, Laux, Hordsworth, Ruberti, Smeekens, Somssich, Weisshaar, Traas
 Bielefeld University, Institute for Genome Research
 Universitaetstrasse 25, D-33594 Bielefeld, Germany
 Email: bernd.weisshaar@uni-bielefeld.de
 AGI: AT2G02450; Seqanalysis: undetermined frame; Translation: no
 full cds detected

Data analysis performed in the frame of REGULATORS (Exploiting
 inter-species conservation in promoter sequences to identify
 regulators of reproductive development and physiological
 performance), a Trilateral Co-Operation in Plant Genomics between
 Spain (MCYT), France (GENOPLANTE) and Germany (GABI) coordinated by
 G. Coupland (coupland-ad-mpiz-koeln.mpg.de). Authors: Vincent
 Thareau (IBP-Orsay UMR8618 CNRS-UPS, thareau-ad-ibp.u-psud.fr) and
 Alain Lecharny (URGV-Evry UMR INRA-CNRS-UEVE,
 lecharny-ad-ibp.u-psud.fr).

Definition of the terms used to describe the quality of the clone:
 The about 2250 sequences from the clone collection were sorted
 according to clones and clustered. If more than one contig was
 formed, the clone was designated 'contamination'. The contigs and
 singletons were blasted against CDS plus pseudogenes from the
 TIGR5 annotation, and the resulting AGI code is presented if more

than 90 percent identity was found. The sequences were also blasted against all TIGR5 introns, and matches longer than 50 bp with 95 percent identity are reported as 'intron found'. The remaining terms for Seqanalysis describe the outcome of the evaluation of the CDS detected after pairwise alignment with CDS plus pseudogenes from the TIGR5 annotation file. The sequences or contigs for which a full CDS with or without STOP codon was detected, a BLASTp against all TIGR5 protein sequences was performed. Full perfect: 100 percent identity; full good: better than 95 percent identity over more than 95 percent of the sequence; partial good: better than 95 percent identity over less than 95 percent of the sequence; weak similarity: less than 95 percent identity over less than 95 percent of the sequence; no similarity: no hit from BLASTp. Note that the collection contains a few clones for which sequencing was not successful, which could have simple technical reasons. Obviously, information about these clones is missing in the submitted data.

Insert Length: 947 Std Error: 0.00
Seq primer: Seta TCGGTTAAGCTAGCATGATCTC.

FEATURES

Location/Qualifiers

1..947

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/ecotype="Columbia"

/db_xref="GABI:1452250"

/db_xref="taxon:3702"

/clone="002-E11"

/lab_host="E. coli DH5alpha"

/clone_lib="MP12-ADIS-065d"

/note="Vector: pDONR201; In the context of the EU-funded Project REGIA (QLG-CT11999-00876, coordinator Javier Paz-Ares), a set of transcription factor ORFs was generated. The ORFs were produced in a decentralized way in the labs of the participants. Most of the ORFs were generated by RT-PCR using cDNA from various A. thaliana tissues as a template. Initially, it was planned to use yeast recombination to move the ORFs from the cloning vectors into target constructs. For this reason, a number of the clones contain 'REGIA tags' (RG_tagl: Spr-AATTCAGCTCACACC-3pr; RG_tag2: Spr-CATGCAATTCGGGGATC-3pr). During the lifetime of the project, the GATEWAY system became available and finally all ORFs were transferred into GATEWAY vectors. At the end of the REGIA project, the clones were collected as plasmid DNA. For end-sequencing at the DNA core facility of the MPI of Plant Breeding Research (ADIS; head: Bernd Weisshaar), the plasmids were re-transformed into DH5alpha, arrayed into 96-well plates and new plasmid DNA was prepared. Re-transformation proved to be difficult because a number of clones displayed poor growth. Data submission has been handled by GabiPD (http://gabi.rzpd.de)."

source

Location/Qualifiers

1..947

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/ecotype="Columbia"

/db_xref="GABI:1452250"

/db_xref="taxon:3702"

/clone="002-E11"

/lab_host="E. coli DH5alpha"

/clone_lib="MP12-ADIS-065d"

/note="Vector: pDONR201; In the context of the EU-funded Project REGIA (QLG-CT11999-00876, coordinator Javier Paz-Ares), a set of transcription factor ORFs was generated. The ORFs were produced in a decentralized way in the labs of the participants. Most of the ORFs were generated by RT-PCR using cDNA from various A. thaliana tissues as a template. Initially, it was planned to use yeast recombination to move the ORFs from the cloning vectors into target constructs. For this reason, a number of the clones contain 'REGIA tags' (RG_tagl: Spr-AATTCAGCTCACACC-3pr; RG_tag2: Spr-CATGCAATTCGGGGATC-3pr). During the lifetime of the project, the GATEWAY system became available and finally all ORFs were transferred into GATEWAY vectors. At the end of the REGIA project, the clones were collected as plasmid DNA. For end-sequencing at the DNA core facility of the MPI of Plant Breeding Research (ADIS; head: Bernd Weisshaar), the plasmids were re-transformed into DH5alpha, arrayed into 96-well plates and new plasmid DNA was prepared. Re-transformation proved to be difficult because a number of clones displayed poor growth. Data submission has been handled by GabiPD (http://gabi.rzpd.de)."

2123 CGAGAGCTCTCCACCGAATATTCGGCGAGCAGCAACAACAGCACAACAG 2182
|||||
829 CGAGAGCTCTCCACCGAATATTCGGCGAGCAGCAGCAACAACAGCACAACAG 888
|||||
2183 TAACTCTGACGTTACCAATGCTGTAGCCATCAAAACATATATGTCCTCAATGCCTTACG 2241
|||||
889 TAACTCTGACGTTACCAATGCTGTAGCCATCAAAACATATATGTCCTCAATGCCTTACG 947
|||||

RESULT 8
AV440745 427 bp mRNA linear EST 18-FEB-2004
LOCUS
DEFINITION
AV440745 Arabidopsis thaliana above-ground organ two to six-week
old Arabidopsis thaliana cDNA clone AP208c08_f 3', mRNA sequence.
ACCESSION
AV440745
VERSION
AV440745.1 GI:7611116
KEYWORDS
EST.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
1 (bases 1 to 427)
Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7 (3), 175-180 (2000)
JOURNAL
PUBMED
10907947
COMMENT
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/
LOCATION/Qualifiers
1..427
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="AP208c08_f"
/tissue_type="aboveground organs"
/dev_stages="two to six-week old"
/clone_lib="Arabidopsis thaliana above-ground organ two to
six-week old"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

ORIGIN
Query Match 11.2%; Score 292.8; DB 1; Length 427;
Best Local Similarity 99.3%; Pred. No. 2.3e-47;
Matches 294; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCAATTGTATCTCTCCACCAAGCATCATCTCCATGAGTAACCAAGTCAACATAAC 60
DB 32 ATGGCAATTGTATCTCTCCACCAAGCATCATCTCCATGAGTAACCAAGTCAACATAAC 91
QY 61 GAAAGAGTATAGAGCAATGATCATAGAGCGCGCCAGAGAGTCAATGTCCTCAATGAA 120
DB 92 GAAAGAGTATAGAGCAATGATCATAGAGCGCGCCAGAGAGTCAATGTCCTCAATGAA 151
QY 121 GATGAAGTGTATGATCATGATCATGATGTCCTCCCGGATTTAGATTCCTATCTACC 180
DB 152 GATGAAGTGTATGATCATGATCATGATGTCCTCCCGGATTTAGATTCCTATCTACC 211
QY 181 GAAGAAGAACTCATAGAGTTTACCTTCGCCGAAAGTTGAAGGCAACGCTTTTATGTA 240
DB 212 GAAGAAGAACTCATAGAGTTTACCTTCGCCGAAAGTTGAAGGCAACGCTTTTATGTA 271
QY 241 GAATCATCATCTTCTCTCGATCTTTATCGCTATGATCTCTGGGAACCTTCCTGGTAA 296
DB 272 GAATCATCATCTTCTCTCGATCTTTATCGCTATGATCTCTGGGAACCTTCCTGGTAA 327

ORIGIN

Query Match 12.1%; Score 314.2; DB 8; Length 947;
Best Local Similarity 92.2%; Pred. No. 1.4e-51;
Matches 331; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 1883 GTATCAAGTCTCTTAATCAATGTCTCATGTATATTAATATATAGGCTGAATATCATTTGGCG 1942
DB 589 GTATCGTCTTCGTACCATGMAACCGAGAAGTACCAAAAGGCTGAAATATCATTTGGCG 648
QY 1943 AGTGTAACAAAGCCAGGAGTAGAAGATCATCTCATGTTACCAAGTCTCTCTCCACAG 2002
DB 649 AGTGTAACAAAGCCAGGAGTAGAAGATCATCTCATGTTACCAAGTCTCTCTCCACAG 708
QY 2003 ACATCATATACCAATCACTCATGATCATCTCCGCTTTAGCTTTAAGACAAACAACACCA 2062
DB 709 ACATCATATACCAATCACTCATGATCATCTCCGCTTTAGCTTTAAGACAAACAACACCA 768
QY 2063 TTTCATCTCTCTTAATCAATTCGACAAACACCTTAACAAACAACATCAACATCT 2122
DB 769 TTTCATCTCTCTTAATCAATTCGACAAACACCTTTAACAACAACATCAACATCT 828

DEFINITION oed90h05.b1 B.oleracea002 Brassica oleracea genomic, genomic survey sequence.

ACCESSION BZ028065

VERSION BZ028065.1 GI:23591339

KEYWORDS GSS.

ORGANISM Brassica oleracea

REFERENCE 1 (bases 1 to 729)
Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T., Nash,W., Rabinowicz,P.D. and Wilson,R.K.
Whole genome shotgun reads from Brassica oleracea unpublished (2002)
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Plate: oed90 row: h column: 05
Seq primer: -21UpOT forward
Class: shotgun
High quality sequence start: 37
High quality sequence stop: 551.

FEATURES
source
1..729
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="B.oleracea002"
/note="Vector: pOTW13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea TO1000DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

ORIGIN
Query Match 8.7%; Score 226.2; DB 9; Length 729;
Best Local Similarity 88.0%; Pred. No. 3.5e-34;
Matches 272; Conservative 0; Mismatches 28; Indels 9; Gaps 2;

QY 1 ATGGCAATTTGATCTCCACACAGCATCTCCATGAGTAAACCAAGTCAACAAATAC 60
|||||
Db 250 ATGGCAATTTGATCTCCACACAGCATCTCCATGAGTAAACCAAGTCAACAAATAC 309
|||||
QY 61 ---GAAAAGGTATAGAACAATGATCATAGAGCGCGCCAAAGAGTCAATGTCCAAAT 117
|||||
Db 310 AGTGATAAGGTGTACAAGAGTGTCTATAGA-----CACGAGAATCATGCCCAAT 363
|||||
QY 118 GAAGATGAAGTGTATGATCATGATCATGATGATGATGATGATGATGATGATGATGAT 177
|||||
Db 364 GATGATGAAGTGTATGATCATGATCATGATGATGATGATGATGATGATGATGATGAT 423
|||||
QY 178 ACCGAGAGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 237
|||||
Db 424 ACAGAGAGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 483
|||||
QY 238 GTAGAACTCATCTTCTCGATCTTATCGTATGATGATGATGATGATGATGATGATGAT 297
|||||
Db 484 GTAGAACTCATCTTCTCGATCTTATCGTATGATGATGATGATGATGATGATGATGAT 543
|||||
QY 298 TATACATTC 306
|||||
Db 544 TATATATTC 552
|||||

RESULT 12
BH999674
LOCUS oef14h08.b1 B.oleracea002 Brassica oleracea genomic, genomic survey sequence.

ACCESSION BH999674

VERSION BH999674.1 GI:23547335

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

REFERENCE 1 (bases 1 to 702)
Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T., Nash,W., Rabinowicz,P.D. and Wilson,R.K.
Whole genome shotgun reads from Brassica oleracea unpublished (2002)
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Plate: oef14 row: h column: 08
Seq primer: -21UpOT forward
Class: shotgun
High quality sequence start: 38
High quality sequence stop: 551.

FEATURES
source
1..702
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="B.oleracea002"
/note="Vector: pOTW13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea TO1000DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

ORIGIN
Query Match 8.4%; Score 219.2; DB 9; Length 702;
Best Local Similarity 81.3%; Pred. No. 8.5e-33;
Matches 309; Conservative 0; Mismatches 53; Indels 18; Gaps 4;

QY 2214 CAAACATATATGTCCTCAATGCTTACGACACAGCAACACATTCGATGCTCTACG 2273
|||||
Db 702 CAAACATATATGTCCTCAATGCTTACGACACAGCAACACATTCGATGCTCTACG 643
|||||
QY 2274 AGAATCATCAAGAGTATGATGAAGTCCATGTTGTCAGATCTTCAAGACTAGTTAA 2333
|||||
Db 642 A-----ATAAGAGAGTATGATGAAGTCCATGTTGTCAGATCTTCAAGACTAGTTAA 589
|||||
QY 2334 TACCAATATCAGATGGAG---GTAACATCAATCACCATATCTTTCAAAATGCTCAACAG 2390
|||||
Db 588 TACCAATATCTGATGGAGTAGTAACATCAATCACCATATCTTTCAAAATGCTCAACAG 529
|||||
QY 2391 TTTTCATCATCTCAACAAACAAATGCTTAAACGAAACGATACAAATGCTGCTCGGGG 2450
|||||
Db 528 TTTTCATCATCTCAACAAACAAATGCTTAAACGAAACGATACAAATGCTGCTCGGGGCA 475
|||||
QY 2451 ACTACAGCGACACGCTTAACTCCCTCAATGCTTAACTCAAGCGGCTTAGCTATGAACATGATTCCT 2510
|||||
Db 474 ACTACAGTG---GGCTTAACTCCCTCAACGCGGCTTAGCTATGAACATGATTCCT 418
|||||
QY 2511 GCAGAAACGATTCCTCAACAAATGCTTAACTCCCTCAAGCGGCTTAGCTATGAACATGATTCCT 2570
|||||
Db 417 GCAGGAGATTCCTCAACAAATGCTTAACTCCCTCAAGCGGCTTAGCTATGAACATGATTCCT 358
|||||
QY 2571 AACAGAGATCATATCTAA 2590
|||||
Db 357 AACAGAGATCATATCTAA 338
|||||

RESULT 13
AY201073/c
LOCUS AY201073

280 bp DNA linear GSS 03-JAN-2005

Search completed: April 6, 2006, 13:42:56
Job time : 9253 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic. - nucleic search, using sw model

Run on: April 6, 2006, 09:19:16 ; Search time 457 Seconds
(without alignments)
10136.380 Million cell updates/sec

Title: US-10-780-703-3

Perfect score: 2606

Sequence: 1 atggcaattgtatctccac.....ctaatattctttaagtaa 2606

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
- 7: /cgn2_6/ptodata/1/ina/pp COMB.seq:*
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
- 9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 97.4 | 3.7 | 1216 | 3 | US-09-533-029-121 |
| 2 | 90.8 | 3.5 | 1212 | 3 | US-09-533-029-3 |
| 3 | 90.6 | 3.5 | 1409 | 3 | US-09-533-029-37 |
| 4 | 88.4 | 3.4 | 914 | 3 | US-09-533-029-23 |
| 5 | 87.8 | 3.4 | 2680 | 3 | US-09-614-408-7 |
| 6 | 86.4 | 3.3 | 1211 | 3 | US-09-614-408-1 |
| 7 | 86.4 | 3.3 | 1211 | 3 | US-09-614-981-1 |
| 8 | 84.8 | 3.3 | 1209 | 3 | US-09-614-408-6 |
| 9 | 83.8 | 3.2 | 187169 | 3 | US-09-949-016-12776 |
| 10 | 83.8 | 3.2 | 191569 | 3 | US-09-949-016-15940 |
| 11 | 81.4 | 3.1 | 1287 | 3 | US-09-889-926-1 |
| 12 | 80.2 | 3.1 | 30820 | 3 | US-09-949-016-17145 |
| 13 | 79.8 | 3.1 | 1141 | 3 | US-09-806-708B-22 |
| 14 | 79.6 | 3.1 | 18773 | 3 | US-09-949-016-14164 |
| 15 | 77.4 | 3.0 | 119153 | 3 | US-09-949-016-12378 |
| 16 | 76.2 | 2.9 | 134987 | 3 | US-09-949-016-15348 |
| 17 | 76.2 | 2.9 | 134987 | 3 | US-09-949-016-15349 |
| 18 | 76.2 | 2.9 | 134987 | 3 | US-09-949-016-15350 |
| 19 | 76.2 | 2.9 | 134987 | 3 | US-09-949-016-15507 |
| 20 | 76.2 | 2.9 | 134987 | 3 | US-09-949-016-15508 |
| 21 | 76.2 | 2.9 | 134987 | 3 | US-09-949-016-15509 |
| 22 | 76 | 2.9 | 913 | 3 | US-09-533-029-73 |
| 23 | 74.2 | 2.8 | 55886 | 3 | US-09-949-016-15129 |
| 24 | 73.6 | 2.8 | 114139 | 3 | US-09-949-016-16536 |

| | | | | | | |
|------|------|-----|--------|---|---------------------|-------------------|
| C 25 | 72.2 | 2.8 | 205044 | 3 | US-09-949-016-15851 | Sequence 15851, A |
| C 26 | 72.2 | 2.8 | 205044 | 3 | US-09-949-016-15852 | Sequence 15852, A |
| C 27 | 72.2 | 2.8 | 205044 | 3 | US-09-949-016-15853 | Sequence 15853, A |
| C 28 | 72.2 | 2.8 | 223471 | 3 | US-09-949-016-12387 | Sequence 12387, A |
| C 29 | 72.2 | 2.8 | 223471 | 3 | US-09-949-016-12724 | Sequence 12724, A |
| C 30 | 72.2 | 2.8 | 223471 | 3 | US-09-949-016-12725 | Sequence 12725, A |
| C 31 | 71.4 | 2.7 | 700 | 3 | US-09-735-271-1038 | Sequence 1038, Ap |
| C 32 | 71.4 | 2.7 | 95255 | 3 | US-09-949-016-17067 | Sequence 17067, A |
| C 33 | 71.4 | 2.7 | 231129 | 3 | US-09-949-016-16110 | Sequence 16110, A |
| C 34 | 71.4 | 2.7 | 262293 | 3 | US-09-949-016-11934 | Sequence 11934, A |
| C 35 | 71 | 2.7 | 18773 | 3 | US-09-949-016-14164 | Sequence 14164, A |
| C 36 | 70.8 | 2.7 | 1141 | 3 | US-09-806-708B-22 | Sequence 22, Appl |
| C 37 | 70.2 | 2.7 | 67755 | 3 | US-09-949-016-17296 | Sequence 17296, A |
| C 38 | 70 | 2.7 | 251672 | 3 | US-09-949-016-11973 | Sequence 11973, A |
| C 39 | 70 | 2.7 | 251682 | 3 | US-09-949-016-13703 | Sequence 13703, A |
| C 40 | 69.4 | 2.7 | 67755 | 3 | US-09-949-002-777 | Sequence 777, App |
| C 41 | 69.2 | 2.7 | 25590 | 3 | US-09-949-016-12900 | Sequence 12900, A |
| C 42 | 69 | 2.6 | 462589 | 3 | US-09-949-016-12412 | Sequence 12412, A |
| C 43 | 69 | 2.6 | 476044 | 3 | US-09-949-016-12386 | Sequence 12386, A |
| C 44 | 68.6 | 2.6 | 263693 | 3 | US-09-949-016-16915 | Sequence 16915, A |
| C 45 | 68.6 | 2.6 | 263694 | 3 | US-09-949-016-16915 | Sequence 16915, A |

ALIGNMENTS

RESULT 1
US-09-533-029-121
; Sequence 121, Application US/09533029
; Patent No. 6664446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddie, James
; APPLICANT: Pineda, Onaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MEI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 121
; LENGTH: 1216
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G501
US-09-533-029-121

| | | | | |
|-----------------------|-------|---|-------|----------------------------------|
| Query Match | 3.7% | Score 97.4; | DB 3; | Length 1216; |
| Best Local Similarity | 64.2% | Pred. No. 1.9e-11; | | |
| Mismatches | 165; | Conservative | 0; | Mismatches 86; Indels 6; Gaps 1; |
| QY | 618 | TGCAGCTATGGCGCGATAGGAGAGAAAGAGTGGTACTTCTATGTGCAAGAGATCGGAA | 677 | |
| Db | 217 | TCCAGAGATGTCTCTACGGAGAGAAAGAGTGGTACTTCTCTACCTAGAGATCGGAA | 276 | |
| QY | 678 | ATATAGAAATGGAGATAGACCGAACCGAGTAACGACTTCAGGATATATGGAAGCCACCG | 737 | |
| Db | 277 | ATACCCAAACGGTTCGCGTCTTAACCGGCGAGCGAACCGGTTATTTGGAAGTACCGG | 336 | |
| QY | 738 | AGCTGATAGGATGATCAGATCGGAGACTTTCGGCTATCGGATTAAGAGAAACCTAGT | 797 | |

Db 337 AGCAGTAAACCGAT-----TGGTAAACCGAAGAGCTTGGGTATCAAGAAAGCACTCGT 390
QY 798 TTTCTACTCTGTAAGCCCTAAAGGCACTCGTACTAGTTGGATCATGAACGAGTATCG 857
Db 391 CTTCTACGAGGAAAGCTCCAAAGGGGATTAAAGCAATTTGGTAATGATGATGATCG 450
QY 858 TCTTCGCGACCATGAAA 874
Db 451 TCTCGCTAATGTTGATA 467

RESULT 2

US-09-533-029-3

; Sequence 3, Application US/09533029

; Patent No. 6664446

; GENERAL INFORMATION:

; APPLICANT: Heard, Jacqueline

; APPLICANT: Broun, Pierre

; APPLICANT: Riechmann, Jose-Luis

; APPLICANT: Keddle, James

; APPLICANT: Pineda, Omaira

; APPLICANT: Adam, Luc

; APPLICANT: Samaha, Raymond

; APPLICANT: Zhang, James

; APPLICANT: Yu, Guo-Liang

; APPLICANT: Ratcliffe, Oliver

; APPLICANT: Pilgrim, Marsha

; APPLICANT: Jiang, Cai-Zhong

; APPLICANT: Reuber, Lynne

; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES

; FILE REFERENCE: MBI-010

; CURRENT APPLICATION NUMBER: US/09/533,029

; CURRENT FILING DATE: 2000-03-22

; EARLIER APPLICATION NUMBER: 60/125,814

; EARLIER FILING DATE: 1999-03-23

; NUMBER OF SEQ ID NOS: 121

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 1212

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; FEATURE:

; OTHER INFORMATION: G759

; US-09-533-029-3

Query Match 3.5%; Score 90.8; DB 3; Length 1212;
Best Local Similarity 61.8%; Pred. No. 4.9e-10;
Matches 162; Conservative 0; Mismatches 97; Indels 3; Gaps 1;

QY 599 AATCATAGATATATGTATATGAGCTATCGCGCGATAGAGAGAAAGTGGTACTTCT 658
Db 165 AATTCGATCCATGGGTTTACCATAAAGCAATTATTTGGAGAAAAAGAAATGGTATTTT 224
QY 659 ATGTGCAAGAGATCGAAATATAGAAATGGAGATAGACCGAAGAGTAAAGCACTTCAG 718
Db 225 TTAGTCCTTAGGATAGAAATATCCAAACGGGTCAAGACTAACCGGTTGCGGATCGG 284
QY 719 GATATTGGAAGCCACCGAGTGTATAGATGATCAGATCGAGACTTCTCGGCCCTATG 778
Db 285 GTTATTGGAAGCTACGGGTACGGATATAAATATC---TCGACGGAAGCAAAAGATTG 341
QY 779 GATTAAAGAAACCCCTAGTTTCTACTCTGGTAAAGCCCTAAAGCACTCTGACTAGTT 838
Db 342 GTATTAAAGAAAGCTTGGTGTGTTTACATCGGAAAGCTCTTAAGGTACTATAAACAATT 401
QY 839 GGATCATGAACAGTATCGTCT 860
Db 402 GGATCATGATGATGCTCT 423

RESULT 3

US-09-533-029-37

; Sequence 37, Application US/09533029
; Patent No. 6664446

; GENERAL INFORMATION:

; APPLICANT: Heard, Jacqueline

; APPLICANT: Broun, Pierre

; APPLICANT: Riechmann, Jose-Luis

; APPLICANT: Keddle, James

; APPLICANT: Pineda, Omaira

; APPLICANT: Adam, Luc

; APPLICANT: Samaha, Raymond

; APPLICANT: Zhang, James

; APPLICANT: Yu, Guo-Liang

; APPLICANT: Ratcliffe, Oliver

; APPLICANT: Pilgrim, Marsha

; APPLICANT: Jiang, Cai-Zhong

; APPLICANT: Reuber, Lynne

; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES

; FILE REFERENCE: MBI-010

; CURRENT APPLICATION NUMBER: US/09/533,029

; CURRENT FILING DATE: 2000-03-22

; EARLIER APPLICATION NUMBER: 60/125,814

; EARLIER FILING DATE: 1999-03-23

; NUMBER OF SEQ ID NOS: 121

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 37

; LENGTH: 1409

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; FEATURE:

; OTHER INFORMATION: G502

; US-09-533-029-37

Query Match 3.5%; Score 90.6; DB 3; Length 1409;
Best Local Similarity 62.7%; Pred. No. 5.6e-10;
Matches 160; Conservative 0; Mismatches 89; Indels 6; Gaps 1;

QY 618 TGCAGCTATGCGCGGATAGAGAGAAAGTGGTACTTCTATGTGCCAAGATCGGAA 677
Db 379 TCCTGGTTTACCTTGTATGTTGAGAGGAAATGGTACTTCTCTCCAGGACAGAAA 438
QY 678 ATATAGAAATGGAGATAGACCGAAGCGAGTAAACGACTTCAGGATATTCGAAAGCCACCG 737
Db 439 ATATCCCAACGGTTGCGGTCCTTAACCGGTCGCTGGTTCTGTTACTGGAAGCTACCG 498
QY 738 AGCTGATAGGATGATCAGATCGGAGACTTCTCGGCTATCGGATTAAGAAACCCCTAGT 797
Db 499 AGCTGATAAACGATCGGA-----CTACCTAAACCGTCGGAATTAAGAAAGCTTGT 552
QY 798 TTTCTACTCTGTTAAAGCCCTAAAGGCACTCGTACTAGTTGGATCATGAACGAGTATCG 857
Db 553 TTTCTACGCGCGCAAGCTCCAAAGGGAGAGAGAAAACCAATTGGATCATGACGAGTACCG 612
QY 858 TCTTCCGCAACATGA 872
Db 613 TCTCGCGGACGTTGA 627

RESULT 4

US-09-533-029-23

; Sequence 23, Application US/09533029

; Patent No. 6664446

; GENERAL INFORMATION:

; APPLICANT: Heard, Jacqueline

; APPLICANT: Broun, Pierre

; APPLICANT: Riechmann, Jose-Luis

; APPLICANT: Keddle, James

; APPLICANT: Pineda, Omaira

; APPLICANT: Adam, Luc

; APPLICANT: Samaha, Raymond

; APPLICANT: Zhang, James

; APPLICANT: Yu, Guo-Liang

; APPLICANT: Ratcliffe, Oliver

; APPLICANT: Pilgrim, Marsha

| | | | |
|----|------|---|------|
| Qy | 737 | GAGCTGATA---GGATGATCAGATCGGAGACTTCTCGGCCTATCGGATTAAGAAGAAACCC | 793 |
| | | | |
| Db | 963 | GAAACCGACAAGCCGGTTTTTACTTCCGCTGGAAACAAAAAGTTGGGGTAAAAAAGGCGC | 1021 |
| Qy | 794 | TAGTTTTCTACTCTGGTAAAGCCCTAAAGGCACTCGTACTAGTTGGATCATGACGAGT | 853 |
| | | | |
| Db | 1023 | TCGTTTTTTACGGCGGTAAACCAACAAAGGGGTAAAAACTAAATTTGGATCGTGCATGAAT | 1081 |
| Qy | 854 | ATCGTCTTCGCAACCATGAAACCGAGAAGTACCAA | 888 |
| | | | |
| Db | 1083 | ACAGAGTTGTAGAAAAATAAAACAAATACACAGCCA | 1117 |

RESULT 6
 US-09-614-408-1
 ; Sequence 1, Application US/09614408
 ; Patent No. 6762347
 ; GENERAL INFORMATION:
 ; APPLICANT: GIOVANNONI, JAMES
 ; APPLICANT: TANKSLEY, STEVEN
 ; APPLICANT: VREBALOV, JULIA
 ; APPLICANT: NOENSIE, FREDERICK
 ; TITLE OF INVENTION: NOR GENE COMPOSITIONS AND METHODS FOR USE THEREOF
 ; FILE REFERENCE: TAMK:213
 ; CURRENT APPLICATION NUMBER: US/09/614,408
 ; CURRENT FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: 60/143,357
 ; PRIOR FILING DATE: 1999-07-12
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 1211
 ; TYPE: DNA
 ; ORGANISM: Tomato
 US-09-614-408-1

RESULT 7
US-09-614-981-1
; Sequence 1, Application US/09614981
; Patent No. 6787687
; GENERAL INFORMATION:
; APPLICANT: GIOVANNONI, JAMES
; APPLICANT: TANKSLEY, STEVEN
; APPLICANT: PADMANABHAN, VEERAPAGAVAN
; APPLICANT: RUEZINSKY, DIANE
; APPLICANT: VREBALOV, JULIA
; APPLICANT: WHITE, RUTH

```

; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 914
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G503
US-09-533-029-23

```

[illegible]

| | Query Match | 3.4% | Score 87.8 | DB 3 | Length 2680 |
|----|-----------------------|---|-------------------|----------|-------------|
| | Best Local Similarity | 60.0% | Pred. No. 2.7e-09 | | |
| | Matches 155 | Conservative 0 | Mismatches 107 | Indels 3 | Gaps 1 |
| QY | 617 | ATCCAGCTATGGCGGCGATAGGAGAGAAAGAGTGGTACTTCTATGTGCGCCAAAGAGATCGGA | 676 | | |
| | | | | | |
| Db | 843 | ACGCAGCTAAGGCAATATTCCGAGAGCAAGATGGTCTCTTTTAGTCCAAAGAGATAGAA | 902 | | |
| | | | | | |
| QY | 677 | AATATAGAAATGGAGATAGACCGAACCCGATACAGACTTCAGGATATTGGAAGGCCACCG | 736 | | |
| | | | | | |
| Db | 903 | AATATCTTAACGGGCGAGGCCAAATCGGCGCTGCAACATCGGCTTATTGGAAGGCTACCG | 952 | | |
| | | | | | |

; TITLE OF INVENTION: RIN GENE COMPOSITIONS AND METHODS FOR USE THEREOF
 ; FILE REFERENCE: TAMK:214
 ; CURRENT APPLICATION NUMBER: US/09/614,981
 ; CURRENT FILING DATE: 2000-07-12
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 1211
 ; TYPE: DNA
 ; ORGANISM: Tomato
 US-09-614-981-1

 Query Match 3.3%; Score 86.4; DB 3; Length 1211;
 Best Local Similarity 59.9%; Pred. No. 4.3e-09;
 Matches 163; Conservative 0; Mismatches 106; Indels 3; Gaps 1;

 QY 620 CAGCTATGCGGGGATAGAGAGAAAGAGTGGTACTTCTTCTATGTCGCAAGAGATCGGAAT 679
 DB 220 CTGCTAAGGCAATATTTCGAGAGCAAGAATGGTCTCTTTTATGTCGCAAGAGATAGAAAAT 279

 QY 680 ATAGAATCGAGATAGACCGGAGTAACCGAGTACGACTTTCAGGATATTCGGAAGCCACCGGAG 739
 DB 280 ATCTAACGGGGGAGCGCAATTCGGGCTGCAACATCGGGTTATTGGAAGGCTACCGGAA 339

 QY 740 CTGATA---GGATGATCAGATCGGAGACTTCTCGGCTATCGGCTATTAAGAAAACCCCTAG 796
 DB 340 CCGACRAGCGGTTTTTACTTCGGTGGACACAAAAGGTTGGGTAAAGGCGCTCG 399

 QY 797 TTTTCTACTCTGTAAAGCCCTAAAGGCATCGTACTAGTTGGATCATGAACGAGTATC 856
 DB 400 TTTTTCACGGCGTAAACCAACCAAGGGGTAAACCTAATTGGATCATGTCATGAATACA 459

 QY 857 GTCTCCGACCATGAACCGGAGAGTACCAA 888
 DB 460 GAGTTGTAGAAAATAAACAATAACAAGCCA 491

RESULT 8
 US-09-614-408-6
 ; Sequence 6, Application US/09614408
 ; Patent No. 6762347
 ; GENERAL INFORMATION:
 ; APPLICANT: GIOVANNONI, JAMES
 ; APPLICANT: TANKSLEY, STEVEN
 ; APPLICANT: VREBALOV, JULIA
 ; APPLICANT: NOENSIE, FREDERICK
 ; TITLE OF INVENTION: NOR GENE COMPOSITIONS AND METHODS FOR USE THEREOF
 ; FILE REFERENCE: TAMK:213
 ; CURRENT APPLICATION NUMBER: US/09/614,408
 ; CURRENT FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: 60/143,357
 ; PRIOR FILING DATE: 1999-07-12
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 1209
 ; TYPE: DNA
 ; ORGANISM: Tomato
 US-09-614-408-6

 Query Match 3.3%; Score 84.8; DB 3; Length 1209;
 Best Local Similarity 59.6%; Pred. No. 9.5e-09;
 Matches 162; Conservative 0; Mismatches 107; Indels 3; Gaps 1;

 QY 620 CAGCTATGCGGGGATAGAGAGAAAGAGTGGTACTTCTTCTATGTCGCAAGAGATCGGAAT 679
 DB 220 CTGGAAGGCAATATTTCGAGAGCAAGAATGGTCTCTTTTATGTCGCAAGAGATAGAAAAT 279

 QY 680 ATAGAATCGAGATAGACCGGAGTAACCGACTTTCAGGATATTCGGAAGCCACCGGAG 739
 DB 280 ATCTAACGGGGGAGCGCAATTCGGGCTGCAACATCGGGTTATTGGAAGGCTACCGGAA 339

 QY 740 CTGATA---GGATGATCAGATCGGAGACTTCTCGGCTATCGGCTATTAAGAAAACCCCTAG 796

DB 340 CCGACAAGCCGGTTTTTACTTCCGGTGGAAACACAAAAGGTTGGGTAATAAAGGCGCTCG 399
 QY 797 TTTTCTACTCTGTAAAGCCCTAAAGGCATCGTACTAGTTGGATCATGAACGAGTATC 856
 DB 400 TTTTTCACGGCGTAAACCAACCAAGGGGTAAACCTAATTGGATCATGTCATGAATACA 459

 QY 857 GTCTCCGACCATGAACCGGAGAGTACCAA 888
 DB 460 GAGTTGTAGAAAATAAACAATAACAAGCCA 491

 RESULT 9
 US-09-949-016-12776/c
 ; Sequence 12776, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12776
 ; LENGTH: 187169
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(187169)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-949-016-12776

Query Match 3.2%; Score 83.8; DB 3; Length 187169;
 Best Local Similarity 49.6%; Pred. No. 5.7e-08;
 Matches 298; Conservative 0; Mismatches 297; Indels 6; Gaps 3;

 QY 1332 TTTAAGTTACCAACATTTTTTTGGAATAATTGGCATATATGAAATATACCAACATATTT 1391
 DB 1325 TTTATTTTATATATATTTTATTTATTTATATATTTATATATTTTATATATATAA 1266

 QY 1392 ATATCGCAACATTTTAAATCTATACGAATGATAACGGTTTATGGAGTAGACCGAAAAAT 1451
 DB 1265 TATATAAATATATATTTTATATATATATATATATATATATATATATATATATATA 1206

 QY 1452 ATTATGATACGGAATAATGACATGAGATAGATAAATACATTTTTTTGGGCTCTTCGACTT 1511
 DB 1205 TATTTA 1146

 QY 1512 ATATGCTGTCACCATTTTGAACCATATAAATTTTCTATGATATATATATATATATAT 1571
 DB 1145 AT 1087

 QY 1572 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1631
 DB 1086 TATATATATATATATTTTATACATATATATATATATATATATATATATATATATAT 1027

 QY 1632 ATTGCACATTCCTGTTACTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1691
 DB 1036 TTTATATTTTATTTTAT 968

 QY 1692 TAAATATATATCTTCTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 1751
 DB 967 ATATATTTTAT 912

; Sequence 17145, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17145
; LENGTH: 30820
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17145

Query Match 3.1%; Score 80.2; DB 3; Length 30820;
Best Local Similarity 47.8%; Pred. No. 2.1e-07;
Matches 361; Conservative 0; Mismatches 383; Indels 11; Gaps 4;
QY 1121 ATGTCATATGCGTGAACCTCTTTAAATGCTGACACATGGCCCATATGTTATAGTAAGTAATC 1180
DB 6830 ATGGGAAACGTATACAATTTGTAAGATACAAGAACACACATCATATTTTAAACACATGTAA 6771
QY 1181 ATTAATAGATAGATGATACATATATATAACACACAGTATCACCTCGACATTCAT 1240
DB 6770 ATTTAT 6711
QY 1241 ATACCTTAATCTCGACAGACATAGTTAGTTCTTACAAATTTATGACATGAATGTTCC 1300
DB 6710 A-----TAATAATATATAAATATATATATATATATATATATATATATATATATATAT 6656
QY 1301 TGCCTCTCCATCAATTAATCAATGCTCTTCTATTTAAAGTTACCCACATTTTGGAAATAA 1360
DB 6655 TTATAATATATATAAAT 6597
QY 1361 TTGGCATATGAATATACCAACATATTTATATCGGAACATTTTAAATCTATACGAAT 1420
DB 6596 TATACAAT 6537
QY 1421 CATACCGGTTTATGGAGTAGACCGAAAAATATATATGTATACGAAAAATGACAAATGGATA 1480
DB 6536 TAT 6477
QY 1481 GATAAATACATTTTGGGCTCTTTCGACTATATATGTCGACCATTTGAAACCAATAAT 1540
DB 6476 ATATATATATCTTATATACAAATATATA-ATATATATATCTTATATACAAATATATAA 6418
QY 1541 TTATAAATTTCTATGTAT 1600
DB 6417 TATTATATCTTATATACAAAT 6358
QY 1601 ACAGGGTTGACATAAT- - -TATCTATGTGTATGTTATTCACATTCACCTTGACTAATAA 1656
DB 6357 ATCTTATATACAAAT 6298
QY 1657 AACTAAATACGCAATTAATAATATAAATAATAAATAATAAATAAATAAATAAATAAATAAATA 1716
DB 6297 TATATACAAATATAAATAATATATCTTATATACAAATATAAATAAATAAATAAATAAATA 6238
QY 1717 TGCATGTACGTCATATAGTAGTACTTAAATTTCTTCTTAAACGTCATCTTTTGTG 1776
DB 6237 ATATAAATAGTAT 6178
QY 1777 CTAATGCTAACTTTACATAGTTTGGAACTCTTCTTCAAAACCATATCTTCGATAAATGA 1836

DB 6177 TATATAAT 6118
QY 1837 TATTTTTCATAGATATTTCTTAGTCTATATATTTGATA 1871
DB 6117 AGTATAAATAGTA 6083
RESULT 13
US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAEI promoters
US-09-806-708B-22

Query Match 3.1%; Score 79.8; DB 3; Length 1141;
Best Local Similarity 11.1%; Pred. No. 1.1e-07;
Matches 114; Conservative 371; Mismatches 538; Indels 0; Gaps 0;
QY 891 GGTATAAATCTACTATAAATCTCTATATATATATATATATATATATATATATATATATATAT 950
DB 85 KKWYBCANNTSBRYHARRWKDMKTAYBMTNKGKTCGRHRYWRWABDVTVDHHYVTA 144
QY 951 AGCTAGGTGGTGGAGCCCTTTTAAATTTGAAATTAATCCCTAGACAGTTTGAATTTTCTT 1010
DB 145 MNNAAWTTMCMMDKDDKRTWMMKKNNATGWDGDDTKYHMMNNNGCBVTVMVRYKTDMD 204
QY 1011 TTTTGACTAGTTTATTTATTTATTTTGGAAATTCGATTAAGATCAAAATACCTGTG 1070
DB 205 SBKRMVGMBSWKKWSYDVTYYWVDDMKCKKRVRRVTRGRMRYVMVABETARRRYN 264
QY 1071 AATGGACTTAAATGTGAGCGCGCTTTCGCTTAAATCCAGAAAAATGTTTCATGTGATG 1130
DB 265 NGWTBAMAYRRTWNNNNNNNAKMKRAKYGMNRABVNSTCTTWKSKTKTKVRTSCWANN 324
QY 1131 CGTGAACCTTTTAAATTTGCTAGACATGGCCCATATGTTATAGTAGAATACATAATAGAT 1190
DB 325 CRAGDANKDHKKWSAAMGYVNNNNNNNTYKKARHBAERWDVWVHSAWKWHAANAAY 384
QY 1191 AGATGCATACACATATATATAACACACACAGATCACACTCGACATTCATATACCTTAAT 1250
DB 385 SRKWTBYKRTWVNNNGTTWKKRMWAWYKMDHDEGTYNNNNNGRTYYGWTNKKGN 444
QY 1251 TCTGCAGACATAGTTAGTTTCTTACAAATTTATGACAAATGTTTCTGCTCTTCTCT 1310
DB 445 WTYKKWANNCKWRADWDHKTCHNNTTWKMTYNNNNYKWSMTNGKSHRBAAYVTWYM 504
QY 1311 CACATTAATTCATGCTCTTCTATTTAAGTTACCAACATTTTGTGAAATAATTTGTCATAT 1370
DB 505 WWMRYAHANNNDYVWKKACTWYKBYVCSKWMNNYAAYTKSWNYTSRYRWRWNNSW 564
QY 1371 ATGAATATACCAACATATTTATGCGAACATTTAAATCTATACGAATGATAACGGTT 1430
DB 565 RWRSDTRSMGRANNYARABHYGYKWNTRWBSHWHBRAGAAYHWMYBACCHMK 624
QY 1431 TATGGAGTAGACCGAAAAATATTTATGTATATCGGAAAAATGCAATGAGTAGATAAATACA 1490

[illegible]

QY 181 GAAGAAGAACTCATAGAGTTTTTACCTTCGCGAAAGATTGAAGCGAAACGCTTTAATGTA 240
Db 181 GAAGAAGAACTCATAGAGTTTTTACCTTCGCGAAAGATTGAAGCGAAACGCTTTAATGTA 240
QY 241 GAACTCATCACTTCTCGATCTTTATCGCTATGATCGCTTGGAACTTCTTGGTAAATAT 300
Db 241 GAACTCATCACTTCTCGATCTTTATCGCTATGATCGCTTGGAACTTCTTGGTAAATAT 300
QY 301 ACATTCACATAAACAACACATAAATCATCTCAAACTATTTGGAAATCTTAATTTCTATCA 360
Db 301 ACATTCACATAAACAACACATAAATCATCTCAAACTATTTGGAAATCTTAATTTCTATCA 360
QY 361 TAGTTTAAGATCTTCTCTCTCTATCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
Db 361 TAGTTTAAGATCTTCT 420
QY 421 ATATATGTACCTTCT 480
Db 421 ATATATGTACCTTCT 480
QY 481 CTTTAAACGCTTCTCCCT 540
Db 481 CTTTAAACGCTTCTCCCT 540
QY 541 TATATATCATAGATATACAAATATGTGTATCTTTTCAATATAGCTTATGTATGTTTAA 600
Db 541 TATATATCATAGATATACAAATATGTGTATCTTTTCAATATAGCTTATGTATGTTTAA 600
QY 601 TCATAGATATATGTATGTACGCTATGCGCGGATAGGAGAGAAAGAGTGGTACTTCTAT 660
Db 601 TCATAGATATATGTATGTACGCTATGCGCGGATAGGAGAGAAAGAGTGGTACTTCTAT 660
QY 661 GTGCCAAGAGATCGGAATATAGAAATGAGATAGACCGAACCGAGTAACGACTTCAGGA 720
Db 661 GTGCCAAGAGATCGGAATATAGAAATGAGATAGACCGAACCGAGTAACGACTTCAGGA 720
QY 721 TATTGGAAGCAACCGAGCTGATAGGATGATCAGATCGAGACTTCTCGGCTATCGGA 780
Db 721 TATTGGAAGCAACCGAGCTGATAGGATGATCAGATCGAGACTTCTCGGCTATCGGA 780
QY 781 TTAAGAAACAACTAGTTTCTACTCTGTTAAAGCCCTTAAAGGCACTCGTACTAGTTGG 840
Db 781 TTAAGAAACAACTAGTTTCTACTCTGTTAAAGCCCTTAAAGGCACTCGTACTAGTTGG 840
QY 841 ATCATGAACGAGTATCGTCTTCGCAACCATGAAACCGAGAGTACAAAGGTATAAAT 900
Db 841 ATCATGAACGAGTATCGTCTTCGCAACCATGAAACCGAGAGTACAAAGGTATAAAT 900
QY 901 CTACTATACCTTATATATATCTTATTCATACATACATAGATATAACCTAGCTAGTGG 960
Db 901 CTACTATACCTTATATATATCTTATTCATACATACATAGATATAACCTAGCTAGTGG 960
QY 961 TGAGGCGCTTTAAATTAATTAATCCCTAGACAGTTTGAATTTTCTTTTGGACTAG 1020
Db 961 TGAGGCGCTTTAAATTAATTAATCCCTAGACAGTTTGAATTTTCTTTTGGACTAG 1020
QY 1021 TTTTATTTATTTATTTTGAATTTGATTCGATTAAGATCAAAATATCTGTGGAATGACTAA 1080
Db 1021 TTTTATTTATTTATTTTGAATTTGATTCGATTAAGATCAAAATATCTGTGGAATGACTAA 1080
QY 1081 ATGTCAGGCGGCTTGGCTTAAATCCAGAAATGTTTCATGTATATGCGTGAACCTCT 1140
Db 1081 ATGTCAGGCGGCTTGGCTTAAATCCAGAAATGTTTCATGTATATGCGTGAACCTCT 1140
QY 1141 TTAATTTGCTAGACATGCGCCATATGTTATAGTAGAATACAATTAATAGATAGATATAC 1200
Db 1141 TTAATTTGCTAGACATGCGCCATATGTTATAGTAGAATACAATTAATAGATAGATATAC 1200
QY 1201 ACATATATAAACAACAAGTATCACTCGACATTTCAATACCTTAAATCTCGAGAGA 1260
Db 1201 ACATATATAAACAACAAGTATCACTCGACATTTCAATACCTTAAATCTCGAGAGA 1260
QY 1261 CATAGTTAGTTTTTCTTACAAATTTATGACATGAATGTTCTCTCTCTCTCAATTAAT 1320

Db 1361 CATAGTTAGTTTTTCTTACAAATTTATGACATGAATGTTCTCTCTCTCTCTCAATTAAT 1320
QY 1321 CATGCTCTCTTATTTAAGTTACCCAAACATTTTTCGAAATTAATTTGGCATATATGATATA 1380
Db 1321 CATGCTCTCTTATTTAAGTTACCCAAACATTTTTCGAAATTAATTTGGCATATATGATATA 1380
QY 1381 CCAACATATTTATATGCGAAACATTTTAAATCTATACGAATGATAAACGCTTTATGGAGTAG 1440
Db 1381 CCAACATATTTATATGCGAAACATTTTAAATCTATACGAATGATAAACGCTTTATGGAGTAG 1440
QY 1441 ACCGAAAAATATTTATGTATACGGAAATGACAAATGAGATAGATAAATACATTTTTTGGGC 1500
Db 1441 ACCGAAAAATATTTATGTATACGGAAATGACAAATGAGATAGATAAATACATTTTTTGGGC 1500
QY 1501 TCTTTCCAGCTTATATGTCGTCACCATTTGAAACCATTAATTTTAAATTTTCTATGATAT 1560
Db 1501 TCTTTCCAGCTTATATGTCGTCACCATTTGAAACCATTAATTTTAAATTTTCTATGATAT 1560
QY 1561 ATATATGATATTTATGATGTATAGACAGCTTAAACACACAGGTTGACATAATTAATC 1620
Db 1561 ATATATGATATTTATGATGTATAGACAGCTTAAACACACAGGTTGACATAATTAATC 1620
QY 1621 TATGCTGATGTATTTGCACATTTCACTTGTACTTAACTAAATTAAGCAATTAATTAAT 1680
Db 1621 TATGCTGATGTATTTGCACATTTCACTTGTACTTAACTAAATTAAGCAATTAATTAAT 1680
QY 1681 ATAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1740
Db 1681 ATAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1740
QY 1741 ACTCTAAATTTCTTAAACGTCATCTTTTGTCTTAACTGCTAACTTTACATAGTTTG 1800
Db 1741 ACTCTAAATTTCTTAAACGTCATCTTTTGTCTTAACTGCTAACTTTACATAGTTTG 1800
QY 1801 TGAATCTTCTTTTCAAAACCATATCTTCGATAAATGATAATTTTTCATAGATATGTTAGTC 1860
Db 1801 TGAATCTTCTTTTCAAAACCATATCTTCGATAAATGATAATTTTTCATAGATATGTTAGTC 1860
QY 1861 TATATTTGATTAATTTGATATATGATCAAGTCTCTTAATCAATGCTCATGTATTAATTAAT 1920
Db 1861 TATATTTGATTAATTTGATATATGATCAAGTCTCTTAATCAATGCTCATGTATTAATTAAT 1920
QY 1921 AGGCTGAAATATCATGTTGCGGAGGTACAAAAGCCAGGAGTAGAAGATCATCCATCGG 1980
Db 1921 AGGCTGAAATATCATGTTGCGGAGGTACAAAAGCCAGGAGTAGAAGATCATCCATCGG 1980
QY 1981 TACCACGTTCTCTCTCCACAGACATCATAACTCACTCATCGACATCATCCGTTTAG 2040
Db 1981 TACCACGTTCTCTCTCCACAGACATCATAACTCACTCATCGACATCATCCGTTTAG 2040
QY 2041 CTTTAAAGCAACAACACCAATTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
Db 2041 CTTTAAAGCAACAACACCAATTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
QY 2101 ACAACAAACATCAAACTCTCGAGAGCTCTCACCGAATATTTCCGCGACGCGCAGCA 2160
Db 2101 ACAACAAACATCAAACTCTCGAGAGCTCTCACCGAATATTTCCGCGACGCGCAGCA 2160
QY 2161 CAACAAACGACCAACCAAGTAACTCTGAGCTTACATTTGCTCTAGCCAATCAAAACA 2220
Db 2161 CAACAAACGACCAACCAAGTAACTCTGAGCTTACATTTGCTCTAGCCAATCAAAACA 2220
QY 2221 TATATGTCGAATGCTTTACGACACAGCAACATTTGATGCTCTAGGAAATC 2280
Db 2221 TATATGTCGAATGCTTTACGACACAGCAACATTTGATGCTCTAGGAAATC 2280
QY 2281 ATCAAGAGGATGATGAATGCTGCTTGTGAGCTCTTCAAGAGACTAGTTAACTACCAA 2340
Db 2281 ATCAAGAGGATGATGAATGCTGCTTGTGAGCTCTTCAAGAGACTAGTTAACTACCAA 2340
QY 2341 TATCAGATGAGGAGTAAACATCAATCAACCAATCTTCAAAATGCTCAACAGTTTCAATCA 2400

Db 2341 TATCAGATGGAGGTAACATCAATCAACCAATATCTTTCAAATTTGCTCAACAGTTTTCATCATA 2400
Qy 2401 CTCACACAAAATGCTTAACGCAACGCAATTAACATTTGGTGGCTGGCGGCACTACAGCGA 2460
Db 2401 CTCACACAAAATGCTTAACGCAACGCAATTAACATTTGGTGGCTGGCGGCACTACAGCGA 2460
Qy 2461 CAACGCTAAATGCTTCAAACTCAAGCGGGCTTAGCTATGAACATGATTCCTCGAGGAACGA 2520
Db 2461 CAACGCTAAATGCTTCAAACTCAAGCGGGCTTAGCTATGAACATGATTCCTCGAGGAACGA 2520
Qy 2521 TTCCAAACAATGCTTTGGGATATGTGGATCAATAGTACCAAGATGGAAGAAACAGAGATC 2580
Db 2521 TTCCAAACAATGCTTTGGGATATGTGGATCAATAGTACCAAGATGGAAGAAACAGAGATC 2580
Qy 2581 ACTATACTAATATCTCTTTAAAGTAA 2606
Db 2581 ACTATACTAATATCTCTTTAAAGTAA 2606

RESULT 2

US-10-225-066A-115

; Sequence 115, Application US/10225066A

; Publication No. US20030226173A1

; GENERAL INFORMATION:

; APPLICANT: Mendel Biotechnology, Inc.

; APPLICANT: RATCLIFFE, Oliver

; APPLICANT: RIECHMANN, Jose Luis

; APPLICANT: ADAM, Luc J

; APPLICANT: DUBELL, Arnold T

; APPLICANT: HEARD, Jacqueline E

; APPLICANT: PILGRIM, Marsha L

; APPLICANT: JIANG, Cai-Zhong

; APPLICANT: REUBER, T. Lynne

; APPLICANT: CREELMAN, Robert A

; APPLICANT: PINEDA, Omalra

; APPLICANT: YU, Guo-Liang

; APPLICANT: BROUN, Pierre E

; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants

; FILE REFERENCE: M810316-2 US

; CURRENT APPLICATION NUMBER: US/10/225,066A

; CURRENT FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: 09/837,444

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: 60/310,847

; PRIOR FILING DATE: 2001-08-09

; PRIOR APPLICATION NUMBER: 60/336,049

; PRIOR FILING DATE: 2001-12-05

; PRIOR APPLICATION NUMBER: 60/338,692

; PRIOR FILING DATE: 2001-12-11

; PRIOR APPLICATION NUMBER: 10/171,468

; PRIOR FILING DATE: 2002-06-14

; NUMBER OF SEQ ID NOS: 1122

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 115

; LENGTH: 1444

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; US-10-225-066A-115

; Query Match 26.4%; Score 687.2; DB 6; Length 1444;

; Best Local Similarity 96.8%; Pred. No. 1e-124;

; Matches 701; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 1883 GTATCAAGTCTCTAATCAATGTGCTCATGTATATATAGGCTGAATATCATTTGTCGCG 1942

Db 669 GTATCGTCTTCGCGACCATGAACCGAGAGTACCAAAAGGCTGAATATCATTTGTCGCG 728

Qy 1943 AGTGTAACAAAGCGCAGAGTAGAAGATCATCCATCGGTACCGATCTCTCTCCACAAG 2002

Db 729 AGTGTAACAAAGCGCAGAGTAGAAGATCATCCATCGGTACCGATCTCTCTCCACAAG 788

Qy 2003 ACATCATAACCAATCAATCGCATCATCCGTTTAGCCTTAGAGACAACCAACACCA 2062

Db 789 ACATCATAACCAATCAATCGCATCATCCGTTTAGCCTTAGAGACAACCAACACCA 848
Qy 2063 TTCTATCTCTCTTAATCAATTCGACACACACCTTTAAACAACAACACATCAACATCT 2122
Db 849 TTCTATCTCTCTTAATCAATTCGACACACACCTTTAAACAACAACACATCAACATCT 908
Qy 2123 CGAAGAAGCTCTCCACCGCAATATTCGCGGACGCGGAGCAGCAACAACAACGACCAACAACAG 2182
Db 909 CGAAGAAGCTCTCCACCGCAATATTCGCGGACGCGGAGCAGCAACAACAACGACCAACAACAG 968
Qy 2183 TAACTCTGAGCTTACCAATGCTCTAGCCAATCAAAAACATATATGCTCAATGCTTACGA 2242
Db 969 TAACTCTGAGCTTACCAATGCTCTAGCCAATCAAAAACATATATGCTCAATGCTTACGA 1028
Qy 2243 CACAAGCAACAACACATGATGCTCTAGGAGAAATCATCAAGACGATGATGAACCTGC 2302
Db 1029 CACAAGCAACAACACATGATGCTCTAGGAGAAATCATCAAGACGATGATGAACCTGC 1088
Qy 2303 CATTGTTGAGCATCTTCAAAAGACTAGTTAACTACCAAAATATCAGATGGAGGTAAACATCAA 2362
Db 1089 CATTGTTGAGCATCTTCAAAAGACTAGTTAACTACCAAAATATCAGATGGAGGTAAACATCAA 1148
Qy 2363 TCACCAATACTTTCAAATGCTCAACAGTTTTCATCATATCTCAACAACAACAATGCTAACGC 2422
Db 1149 TCACCAATACTTTCAAATGCTCAACAGTTTTCATCATATCTCAACAACAACAATGCTAACGC 1208
Qy 2423 AAACGCAATTAACAATGCTGGCTGGCGGACTACAGGACGACGCTAATGCTCAAACTCA 2482
Db 1209 AAACGCAATTAACAATGCTGGCTGGCGGACTACAGGACGACGCTAATGCTCAAACTCA 1268
Qy 2483 AGCGGCTTAGCTATGAACATGATTCCTGCAGGAACGATTCCTCAAAACATGCTTTGTGGGA 2542
Db 1269 AGCGGCTTAGCTATGAACATGATTCCTGCAGGAACGATTCCTCAAAACATGCTTTGTGGGA 1328
Qy 2543 TATGTGGAATCCAATAGTACCAAGTGAAGAAACAGAGATCACTATCTAATATTCCTTTAA 2602
Db 1329 TATGTGGAATCCAATAGTACCAAGTGAAGAAACAGAGATCACTATCTAATATTCCTTTAA 1388
Qy 2603 GTAA 2606
Db 1389 GTAA 1392

RESULT 3

US-10-225-067-87

; Sequence 87, Application US/10225067

; Publication No. US20040019925A1

; GENERAL INFORMATION:

; APPLICANT: Mendel Biotechnology, Inc.

; APPLICANT: Heard, Jacqueline E.

; APPLICANT: Riechmann, Jose Luis

; APPLICANT: Creelman, Robert A.

; APPLICANT: Keddle, James

; APPLICANT: Pilgrim, Marsha L.

; APPLICANT: Dubell, Arnold T.

; APPLICANT: Jiang, Cai-Zhong

; APPLICANT: Ratcliffe, Oliver

; APPLICANT: Pineda, Omalra

; APPLICANT: Yu, Guo-Liang

; APPLICANT: Broun, Pierre E.

; TITLE OF INVENTION: BIOCHEMISTRY-RELATED POLYNUCLEOTIDES AND
; POLYPEPTIDES IN PLANTS

; FILE REFERENCE: 51442002042

; CURRENT APPLICATION NUMBER: US/10/225,067

; CURRENT FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: 60/310,847

; PRIOR FILING DATE: 2001-08-09

; PRIOR APPLICATION NUMBER: 60/336,049

; PRIOR FILING DATE: 2001-11-19

; PRIOR APPLICATION NUMBER: 60/338,692

; PRIOR FILING DATE: 2001-12-11

; PRIOR APPLICATION NUMBER: 10/171,468

; PRIOR FILING DATE: 2002-06-14

```

; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 87
; LENGTH: 1444
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (148)...(1392)
US-10-225-067-87

Query Match
Best Local Similarity 26.4%; Score 687.2; DB 7; Length 1444;
Matches 701; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1893 GTATCAAGTCTCTAATCAATGTGCTCATGTATATATTAGGCTGAATATCATTTGCGG 1942
Db 669 GTATCGTCTTCGGACCACTGAACCGGAGATGACCAAAAGGCTGAATATCATTTGCGG 728

QY 1943 AGTGTAACAAAGGCGAGGAGTAGAGATCATCCATCGGTACCACGTTCTCTCCACAA 2002
Db 729 AGTGTAACAAAGGCGAGGAGTAGAGATCATCCATCGGTACCACGTTCTCTCCACAA 788

QY 2003 ACATCATTAACCAATCACTCATCGACATCACTCCGTTTAGCCTTAAGCAACAACACCA 2062
Db 789 ACATCATTAACCAATCACTCATCGACATCACTCCGTTTAGCCTTAAGCAACAACACCA 848

QY 2063 TTATCTCTCTCTAATCAATTCGCAACAACCTTAACAACAACAACAACAACAACAATCT 2122
Db 849 TTATCTCTCTCTAATCAATTCGCAACAACCTTAACAACAACAACAACAACAACAATCT 908

QY 2123 CGAAGAGCTCTCCACCGAATATTCGGGACGCGGAGCAGCAACAACAACGACCAACAAG 2182
Db 909 CGAAGAGCTCTCCACCGAATATTCGGGACGCGGAGCAGCAACAACAACGACCAACAAG 968

QY 2183 TAACCTCTGACGTTACATTTGCTTAGCAATCAAAACATATATCTGCAATGCTTACGA 2242
Db 969 TAACCTCTGACGTTACATTTGCTTAGCAATCAAAACATATATCTGCAATGCTTACGA 1028

QY 2243 CACAAGCAACAACACATGATGATCTCTACGAGAAATCATCAAGACGATGATGAACATGC 2302
Db 1029 CACAAGCAACAACACATGATGATCTCTACGAGAAATCATCAAGACGATGATGAACATGC 1088

QY 2303 CATTTGACGATCTTCAAGACTAGTTAACTACCAATATCATGAGGTAAACATCAA 2362
Db 1089 CATTTGACGATCTTCAAGACTAGTTAACTACCAATATCATGAGGTAAACATCAA 1148

QY 2363 TCACCAATATCTTCAAAATGCTCAACAGTTTTCATCATCTCAACAACAATAATGCTAA 2422
Db 1149 TCACCAATATCTTCAAAATGCTCAACAGTTTTCATCATCTCAACAACAATAATGCTAA 1208

QY 2423 AAACGCAATTAATTTGGTGGCTGCGGAGTACAGCAACAAGCTTAATGCTTCAAACTCA 2482
Db 1209 AAACGCAATTAATTTGGTGGCTGCGGAGTACAGCAACAAGCTTAATGCTTCAAACTCA 1268

QY 2483 AGCGGGTTAGTATGAACATGATTTCTCGAGAGAGGATTCACAATGCTTTGTTGGA 2542
Db 1269 AGCGGGTTAGTATGAACATGATTTCTCGAGAGAGGATTCACAATGCTTTGTTGGA 1328

QY 2543 TATGTGGATCCAATAGTACAGATGGAACAGAGATCACTACTAATATTTCTTTTAA 2602
Db 1329 TATGTGGATCCAATAGTACAGATGGAACAGAGATCACTACTAATATTTCTTTTAA 1388

QY 2603 GTAA 2606
Db 1389 GTAA 1392

```

```

RESULT 4
US-10-374-780A-2475
; Sequence 2475, Application US/10374780A
; Publication No. US2004001927A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James
; APPLICANT: Broun, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2475
; LENGTH: 1444
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G962
US-10-374-780A-2475

```

```

Query Match 26.4%; Score 687.2; DB 7; Length 1444;
Best Local Similarity 96.8%; Pred. No. 1e-124;
Matches 701; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1893 GTATCAAGTCTCTAATCAATGTGCTCATGTATATATTAGGCTGAATATCATTTGCGG 1942
Db 669 GTATCGTCTTCGGACCACTGAACCGGAGATGACCAAAAGGCTGAATATCATTTGCGG 728

QY 1943 AGTGTAACAAAGGCGAGGAGTAGAGATCATCCATCGGTACCACGTTCTCTCCACAA 2002
Db 729 AGTGTAACAAAGGCGAGGAGTAGAGATCATCCATCGGTACCACGTTCTCTCCACAA 788

QY 2003 ACATCATTAACCAATCACTCATCGACATCACTCCGTTTAGCCTTAAGCAACAACACCA 2062
Db 789 ACATCATTAACCAATCACTCATCGACATCACTCCGTTTAGCCTTAAGCAACAACACCA 848

QY 2063 TTATCTCTCTCTAATCAATTCGCAACAACCTTAACAACAACAACAACAACAATCT 2122
Db 849 TTATCTCTCTCTAATCAATTCGCAACAACCTTAACAACAACAACAACAACAATCT 908

QY 2123 CGAAGAGCTCTCCACCGAATATTCGGGACGCGGAGCAGCAACAACAACGACCAACAAG 2182
Db 909 CGAAGAGCTCTCCACCGAATATTCGGGACGCGGAGCAGCAACAACAACGACCAACAAG 968

QY 2183 TAACCTCTGACGTTACATTTGCTTAGCAATCAAAACATATATCTGCAATGCTTACGA 2242

```

Db 969 TAACTCTGACGTTACCAATGCTCTAGCCAAATCAAAACATATATCGTCCAATGCTTACGA 1028
 Qy 2243 CACAAGCAACACATGATGATGCTCTAGCGAATCATCAAGAAATCATCAAGAGATGATGAACCTGC 2302
 Db 1029 CACAAGCAACACATGATGATGCTCTAGCGAATCATCAAGAAATCATCAAGAGATGATGAACCTGC 1088
 Qy 2303 CATTGTTGACGATCTTCAAGAGCTAGTTAACTTACCAATATCAGATGAGAGTAAACATCAA 2362
 Db 1089 CATTGTTGACGATCTTCAAGAGCTAGTTAACTTACCAATATCAGATGAGAGTAAACATCAA 1148
 Qy 2363 TCACCAATATCTTCAAAATGCTCAACAGTTTCATCATCTCACTCACTCACTCACTCACTCACT 2422
 Db 1149 TCACCAATATCTTCAAAATGCTCAACAGTTTCATCATCTCACTCACTCACTCACTCACTCACT 1208
 Qy 2423 AAACGCTTACAAATGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 2482
 Db 1209 AAACGCTTACAAATGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 1268
 Qy 2483 AGCGGCTTACGATGATGAACATGATTCCTGCGGAGCAAGATTCCTGCGGAGCAAGATTCCTGCGG 2542
 Db 1269 AGCGGCTTACGATGATGAACATGATTCCTGCGGAGCAAGATTCCTGCGGAGCAAGATTCCTGCGG 1328
 Qy 2543 TATGTGGAATCCAATAGTACCAGATGGAACAGAGATCACTATATATTCCTTTTAA 2602
 Db 1329 TATGTGGAATCCAATAGTACCAGATGGAACAGAGATCACTATATATTCCTTTTAA 1388
 Qy 2603 GTAA 2606
 Db 1389 GTAA 1392

RESULT 5

US-10-225-066A-115
 ; Sequence 115, Application US/10225066A
 ; Publication No. US20050160493A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Mendel Biotechnology, Inc.
 ; APPLICANT: RATCLIFFE, Oliver
 ; APPLICANT: RIECHMANN, Jose Luis
 ; APPLICANT: ADAM, Luc J
 ; APPLICANT: DUBELL, Arnold T
 ; APPLICANT: HEARD, Jacqueline E
 ; APPLICANT: PILGRIM, Marsha L
 ; APPLICANT: JIANG, Cai-Zhong
 ; APPLICANT: REUBER, T. Lynne
 ; APPLICANT: CREELMAN, Robert A
 ; APPLICANT: YU, Guo-Liang
 ; APPLICANT: BROWN, Pierre E
 ; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
 ; FILE REFERENCE: MBI0036-2 US
 ; CURRENT APPLICATION NUMBER: US/10/225,066A
 ; CURRENT FILING DATE: 2002-08-09
 ; PRIOR APPLICATION NUMBER: 09/837,444
 ; PRIOR FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: 60/310,847
 ; PRIOR FILING DATE: 2001-08-09
 ; PRIOR APPLICATION NUMBER: 60/336,049
 ; PRIOR FILING DATE: 2001-12-05
 ; PRIOR APPLICATION NUMBER: 60/338,692
 ; PRIOR FILING DATE: 2001-12-11
 ; PRIOR APPLICATION NUMBER: 10/171,468
 ; PRIOR FILING DATE: 2002-06-14
 ; NUMBER OF SEQ ID NOS: 1122
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 115
 ; LENGTH: 1444
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; US-10-225-066A-115
 Query Match 26.4%; Score 687.2; DB 9; Length 1444;
 Best Local Similarity 96.8%; Pred. No. 1e-124;

Matches 701; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
 Qy 1883 GTATCAAGTCTCTAATCAATGCTCATGTATATATATAGCTGAATATCATTTGTGCGG 1942
 Db 669 GTATCGCTTTCCGACCATGAAGAGGAGTACCAAAAGGCTGAATATCATTTGTGCGG 728
 Qy 1943 AGTGTACAAAAGGCGAGGAGTAGAAGATCATCCATCGGTACGATCATCGGTCTCTCCACAAG 2002
 Db 729 AGTGTACAAAAGGCGAGGAGTAGAAGATCATCCATCGGTACGATCATCGGTCTCTCCACAAG 788
 Qy 2003 ACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 2062
 Db 789 ACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 848
 Qy 2063 TTCACTCTCTCTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 2122
 Db 849 TTCACTCTCTCTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 908
 Qy 2123 CGAGAAGCTCTCCACCGAATATTCGGGCGAGCGGAGCAACAAACAAACAAACAAACAAACAG 2182
 Db 909 CGAGAAGCTCTCCACCGAATATTCGGGCGAGCGGAGCAACAAACAAACAAACAAACAAACAG 968
 Qy 2183 TAATCTGAGCTTACCAATGCTCTAGCAATCAAAACATATATCGTCAATGCTTACGA 2242
 Db 969 TAATCTGAGCTTACCAATGCTCTAGCAATCAAAACATATATCGTCAATGCTTACGA 1028
 Qy 2243 CACAAGCAACACATGATGATGCTCTAGCAATCAATCAATCAATCAATCAATCAATCAATCAAT 2302
 Db 1029 CACAAGCAACACATGATGATGCTCTAGCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1088
 Qy 2303 CATTGTTGACGATCTTCAAAAGAGTAGTTAACTTACCAATATCATCAAGAGTAAACATCAA 2362
 Db 1089 CATTGTTGACGATCTTCAAAAGAGTAGTTAACTTACCAATATCATCAAGAGTAAACATCAA 1148
 Qy 2363 TCACCAATATCTTCAAAATGCTCTCAAGATTCATCATATCAACAAACAAACAAACAAACAGC 2422
 Db 1149 TCACCAATATCTTCAAAATGCTCTCAAGATTCATCATATCAACAAACAAACAAACAAACAGC 1208
 Qy 2423 AAACGCTTACAAATGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 2482
 Db 1209 AAACGCTTACAAATGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 1268
 Qy 2483 AGCGGCTTACGATGATGAACATGATTCCTGCGGAGCAAGATTCCTGCGGAGCAAGATTCCTGCGG 2542
 Db 1269 AGCGGCTTACGATGATGAACATGATTCCTGCGGAGCAAGATTCCTGCGGAGCAAGATTCCTGCGG 1328
 Qy 2543 TATGTGGAATCCAATAGTACCAGATGGAACAGAGATCACTATATATTCCTTTTAA 2602
 Db 1329 TATGTGGAATCCAATAGTACCAGATGGAACAGAGATCACTATATATTCCTTTTAA 1388
 Qy 2603 GTAA 2606
 Db 1389 GTAA 1392

RESULT 6

US-10-780-703-1
 ; Sequence 1, Application US/10780703
 ; Publication No. US20050034194A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lee, Jong Seob
 ; APPLICANT: Kim, Yun Hee
 ; APPLICANT: Choi, Eun Kyung
 ; APPLICANT: Yoo, So Yeon
 ; APPLICANT: Ahn, Ji Hoon
 ; APPLICANT: Choi, Yang Do
 ; TITLE OF INVENTION: Gene Controlling Flowering Time of Plants and Method for
 ; FILE REFERENCE: 012679-105
 ; CURRENT APPLICATION NUMBER: US/10/780,703
 ; CURRENT FILING DATE: 2004-02-19
 ; PRIOR APPLICATION NUMBER: KR 10-2003-10772
 ; PRIOR FILING DATE: 2003-02-20

; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1140
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)....(1137)
 ; OTHER INFORMATION: CDNA of LOV1 gene
 US-10-780-703-1

Query Match 17.9%; Score 467.2; DB 8; Length 1140;
 Best Local Similarity 82.3%; Pred. No. 1.6e-81;
 Matches 596; Conservative 0; Mismatches 23; Indels 105; Gaps 1;
 QY 1883 GTATCAAGTCTCTTAATCAATGTCTCATGTATAATATATAGTGTGAATATCATTTGTGCG 1942
 DB 522 GTATCGTCTTCGCGACCATGAAACCGAGAGTACCAAAAGGCTGAAATATCATTTGTGCG 581
 QY 1943 AGTGTACAAAAGGCCAGGAGTAGAAGATCATTCATCGGTACCAAGTTCTCTCTCCACAG 2002
 DB 582 AGTGTACAAAAGGCCAGGAGTAGAAGATCATTCATCGGTACCAAGTTCTCTCTCCACAG 641
 QY 2003 ACATCATACCAATACTCATCGCATCTCCGTTTGTAGCTTAAAGCAACAACACCA 2062
 DB 642 ACATCATACCAATACTCATCGCATCTCCGTTTGTAGCTTAAAGCAACAACACCA 701
 QY 2063 TTCTATCTCTCTTAATCAATTCGCAACAACCTTAAACAACAACAACATCAACAATCT 2122
 DB 702 TTCTATCTCTCTTAATCAATTCGCAACAACCTTAAACAACAACAACATCAACAATCT 761
 QY 2123 CGAAGAGCTCTCCACCGAATATTCGCGGACGCGCAGCAACAACAACACCAACAAG 2182
 DB 762 CGAAGAGCTCTCCACCGAATATTCGCGGACGCGCAGCAACAACAACACCAACAAG 821
 QY 2183 TAACTCTGAGCTTACCATTCGCTAGCCTAAATCAAAATATATCGTCCAAATGCTTACGA 2242
 DB 822 TAACTCTGAGCTTACCATTCGCTAGCCTAAATCAAAATATATCGTCCAAATGCTTACGA 881
 QY 2243 CACAAGCAACAACAATGATGTCTCTACGAGAAATCATCAAGACGATGATGAAACTGC 2302
 DB 882 CACAAGCAACAACAATGATGTCTCTACGAGAAATCATCAAGACGATGATGAAACTGC 941
 QY 2303 CATTTGTGAGCTTCTCAAGACTGTTAACTACCAAAATATCATGATGGAGGTAAATCAA 2362
 DB 942 CATTTGTGAGCTTCTCAAGACTGTTAACTACCAAAATATCATGATGGAGGTAAATCAA 989
 QY 2363 TCACCAATACTTTCAAATTTGCTCAACAGTTTCATCATCTCAACAACAATAATGCTAACGC 2422
 DB 990 ----- 989
 QY 2423 AAACGATTACAAATTTGGTGGCTGCGGCGACTACAGCGCAACGCTAATGCTCAAACTCA 2482
 DB 990 -----AGCGACAACGCTAATGCTCAAACTCA 1016
 QY 2483 AGCGGCTTAGCTATGAACATGATTCCTCGAGGAACGATTCACAAATGCTTTGTGGA 2542
 DB 1017 AGCGGCTTAGCTATGAACATGATTCCTCGAGGAACGATTCACAAATGCTTTGTGGA 1076
 QY 2543 TATGTGAATCCATAGTACCATGGAACAGAGATCACTACTAATATTCCTTTAA 2602
 DB 1077 TATGTGAATCCATAGTACCATGGAACAGAGATCACTACTAATATTCCTTTAA 1136
 QY 2603 GTAA 2606
 DB 1137 GTAA 1140

RESULT 7
 US-10-424-599-121054
 ; Sequence 121054, Application US/10424599
 ; Publication No. US20040031072A1

; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 121054
 ; LENGTH: 785
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_80317C.1
 US-10-424-599-121054

Query Match 7.1%; Score 184.4; DB 7; Length 785;
 Best Local Similarity 78.4%; Pred. No. 5.2e-26;
 Matches 221; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
 QY 618 TCCAGCTATGGCGCGATAGGAGAGAGTGGTACTTCTATGTGCCAAGAGATCGGAA 677
 DB 358 TCCGCGCTTGGCAGCTATTGGTGAGAGAGTGGTACTTCTATGTGCCAAGAGATCGGAA 417
 QY 678 ATATAGAAATGGAGATAGACCGAACCGAGTAACGATTCAGGATATTCGAAAGCCACCGG 737
 DB 418 GTATCGCAACGGTATCGTCCCATCGTGTACTACCTCTGGGTATTTGGAAGCAACAGG 477
 QY 738 AGCTGATAGGATGATCAGATCGGAGACTTCTCGGCTATCGGATTAAGAAACCCCTACT 797
 DB 478 AGCTGATAGGATGATCAGGACGAGAAATTTTCGCTCTATCGCCTCAAGAAACCCCTACT 537
 QY 798 TTTCTACTCTGTAAAGCCCTAAAGGCACTCGTACTAGTTGGATCATGAACGAGTATCG 857
 DB 538 TTTCTATTCTGGAAAGCTCTTAAAGGCATACGAACCACTGGATTATGAACGAGTATCG 597
 QY 858 TCTTCGCGACCATGAAACCGGAGAGTACCAAAAGGTATAAAT 899
 DB 598 CTTGCGCAACATGAAACTGAACGATATATCAAAAGGCTGAGAT 639

RESULT 8
 US-10-437-963-88434
 ; Sequence 88434, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 88434
 ; LENGTH: 1433
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_87286C.1
 US-10-437-963-88434

Query Match 6.3%; Score 164; DB 7; Length 1433;
 Best Local Similarity 70.8%; Pred. No. 6.9e-22;

[illegible]

```

RESULT 9
US-10-437-963-81769
; Sequence 81769, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 81769
; LENGTH: 743
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(743)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_81262C.1
US-10-437-963-81769

```

| | Query Match | 6.0%; | Score 155.6; | DB 7; | Length 743; |
|----|-----------------------|---|--------------------|----------------|-------------------|
| | Best Local Similarity | 71.0%; | Pred. No. 2.3e-20; | | |
| | Matches 206; | Conservative | 0; | Mismatches 84; | Indels 0; Gaps 0; |
| QY | 618 | TCGAGCTATCGCGCGATAGAGNAGAAAGATGGTACTTCTTATGTGCCAAGAGATCGAA | 677 | | |
| Db | 336 | TCCCGCTCTGGCCCTCAATTGGGGACAGGAGTGGTTCTTCTATGTGCCAAGGGACCGCAA | 395 | | |
| QY | 678 | ATATAGAAATGGAGATAGACCGAACCCAGATAAACGACTTTCAGGATATTGGAAAAAGCCACCGG | 737 | | |
| Db | 396 | GTACCGGAAACGGCGATCGGCCCAACCCAGTCAGCCCATCGGGGTACTGGGAAGGCCACAGG | 455 | | |
| QY | 738 | AGCTGATAGGATGATCAGATCGGAGACTTCTTCGCCCTATCGGATTTAAAGAAAAACCCCTAGT | 797 | | |
| Db | 456 | GGCGGATAGGATGGTGAAGGTGGAAGGTGACCGGTCTTATCGGCTCTCAAGAAGACGCTTCGT | 515 | | |

| | | | |
|----|-----|---|-----|
| Qy | 798 | TTTCTACTCTGGTAAAGCCCTAAAGGCACTCGTACTAGTTGGATCATGAAACGAGTATCG | 857 |
| Db | 516 | CTTCTCTGTCGCAAGGCGGCCAAGGGGCTTCGCAGCAGCTGGAATCATGAAACGAGTATCG | 575 |
| Qy | 858 | TCCTCCGCACCATGAAACCGAGAGTACCAAAGGTATAAATTTCTACTAT | 907 |
| Db | 576 | TCTCCCTCATGGCGACGTGACCGCTACCAAAGGTAGAAATTTTCGCTGT | 625 |

RESULT 10
 US-10-425-115-101230
 ; Sequence 101230, Application US/10425115
 ; Publication No. US20040214272A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369325
 ; SEQ ID NO 101230
 ; LENGTH: 390
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(390)
 ; OTHER INFORMATION: unsure at all n locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: MRT4577_23835C.1
 US-10-425-115-101230

| Query Match | 5.6% | Score 145.2; | DB 8; | Length 390; |
|-----------------------|-----------------|---|-----------|-------------|
| Best Local Similarity | 69.6% | Pred. No. 1.9e-18; | | |
| Matches 195; | Conservative 0; | Mismatches 85; | Indels 0; | Gaps 0; |
| Qy | 620 | CACGTATGGCGCGATAGGAGAGAAAGAGTGGTACTTCTATGTGCCAAGAGATCGGAAT | 679 | |
| | | | | |
| Db | 55 | CATCAATGGCCGCGTGGGGAGAAAGAGGGGTAGTTCTACGTCCCGAGGGACCGCAAGT | 114 | |
| | | | | |
| Qy | 680 | ATAGAAATGGAGATAGACCGAACCGAGTACGACTTCAGGATATTGGAAAGCCACCGGAG | 739 | |
| | | | | |
| Db | 115 | ACCGCAACGCGAGACGGCCACCGGGTGACGGGTCTGGGCTACTGGAAGGCCACGGCGC | 174 | |
| | | | | |
| Qy | 740 | CTGATAGGATGATCAGATCGGAGACTTCTCGGGCTATCGGATTAAGAAAAACCCTAGTTT | 799 | |
| | | | | |
| Db | 175 | CCGACCGCACCATCAGGCGGACCAACGCGCGCCATCGGGCTCAAGAACGCTCGTCT | 234 | |
| | | | | |
| Qy | 800 | TCTACTCTGGTAAAGCCCTTAAAGGCACCTCGTACTAGTTGGATCATGAACGAGTATCGTC | 859 | |
| | | | | |
| Db | 235 | TCTACTCCGCAAGGCGCCCAAGGGCGTNCGACGAGCTGGATCATGAACGAGTACCGCC | 294 | |
| | | | | |
| Qy | 860 | TTCCGCACCATGAACCGGAGAGTACCAAAAGGTATAAAT | 899 | |
| | | | | |
| Db | 295 | TCCGCGCCGACGACACCGACCGTTTACCAAGAACCGGAAT | 334 | |
| | | | | |

RESULT 11
US-10-425-115-173486
; Sequence 173486, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants

FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 173486
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_897C.1
US-10-425-115-173486

Query Match 5.4%; Score 140; DB 8; Length 792;
Best Local Similarity 69.2%; Pred. No. 2.8e-17;
Matches 191; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
QY 618 TCAGCTATGGCGGCGATAGGAGAGAAAGAGTGGTACTTCTATGTGCCAAGAGATCGGAA 677
DB 417 TCCAGCAATGGCGGTGATGGCGGGAAGAGTGGTCTTCTACGTGCGGAGGACCGTAA 476
QY 678 ATATAGAAATGAGATAGACCGAACCAGTAAAGACTTCAGGATATTTGAAAGCCACCGG 737
DB 477 GTACCGGAACGAGACCGCCCAACCGGTGACGGCTCGGGGTACTGGAAGCGACCGG 536
QY 738 AGCTGATAGATGATCAGATCGGAGACTTCTCGGCTATCGGATTAAGAAACCTAGT 797
DB 537 CCGTACCGGATGATCCGAGGCGAGAACACCGCCCATCGGCTTGAGAGACGCTCGT 596
QY 798 TTCTACTCTGTAAGCCCTTAAAGGCACTCGTACTAGTTGGATCATGAACGAGTATCG 857
DB 597 GTTCTACTCCGCAAGGCGCCCAAGGGCTCGCAGCAGCTGATCATGAACGAGTACCG 656
QY 858 TTTCCGCAACATGAACACCGAGAAGTACCAAAAGT 893
DB 657 CTTCCGCGCACCAACCGACCGCGCCGACCCATTGAT 692

RESULT 12
US-10-425-115-10100
; Sequence 10100, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 10100
; LENGTH: 1291
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_109212C.1
US-10-425-115-10100

Query Match 5.0%; Score 131.2; DB 8; Length 1291;
Best Local Similarity 64.5%; Pred. No. 1.8e-15;
Matches 196; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
QY 620 CAGCTATGGCGCGATAGGAGAGAAAGTGGTACTTCTATGTGCCAAGAGATCGGAAT 679
DB 286 CTGCTCTGGCTTGGTACGTCGGGACAGAGTGGTACTTCTACGTGCGGAGACCGCACT 345
QY 680 ATAGAAATGAGATAGACCGAACCAGTAAAGACTTCAGGATATTTGAAAGCCACCGGAG 739
DB 346 ACCGCAACGGGACCGCCCAACCGGTTCACGGCTACGCGCTACTGGAAGCGACGGGG 405
QY 740 CTGATAGGATGATCAGATCGGAGACTTCTCGGCTATCGGATTAAGAAACCTTAGTT 799

DB 406 CCGACCGCACGGTCTACGTGAGGTCAAGCGCGCGCTCAAGAAAGCGCTGTGT 465
QY 800 TCTACTCTGGTAAAGCCCTTAAAGCACTCTGACTAGTTGGATCATGAACGAGTATCGTC 859
DB 466 TCTACGTGGGAGAGCGGCCCAAGGACTTCAGGAGCAGCTGATCATGAACGAGTACCGCC 525
QY 860 TTCCGCACCATGAACCGAGAAGTACCAAAAGGTATAAATTTCTACTATACTCTATATAT 919
DB 526 TTCCTTTCAGGCGGCGCGCCGCTACCAAAAGGAATTTCTTTTGGCAAGGTGATCAAAAC 585
QY 920 ATCC 923
DB 586 GACC 589

RESULT 13
US-10-424-599-52441
; Sequence 52441, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 52441
; LENGTH: 1640
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_18368C.1
US-10-424-599-52441

Query Match 4.4%; Score 115; DB 7; Length 1640;
Best Local Similarity 68.0%; Pred. No. 3.1e-12;
Matches 176; Conservative 0; Mismatches 80; Indels 3; Gaps 1;

QY 609 ATATGTATATGCACTATGGCGGCGATAGGAGAGAAAGTGGTACTTCTATGTGCCAAG 668
DB 286 ATGGTCTTCCAGTATAGGCGCGTTTGGAGAGAGAGTGGTACTTCTCAGTCCAAAG 345
QY 669 AGATCGGAAATATAGAAATGGAGATAGACCGAACCAGTAACGACTTCAGGATATTTGAA 728
DB 346 AGACAGGAAGTACCGGATGTTTCCAGCAACAAACAGAGTTGCGGGTTCTGGGTATTTGAA 405
QY 729 AGCACCGGAGCTGATAGGATGATCAGATCGGAGACTTCTCGGCTATCGGATTAAGAA 788
DB 406 AGCCACTGGAATGACAAATATCATCCACTGA---AGGTAGAAAAGTTGGCATAAAAA 462
QY 789 AACCTAGTTTCTACTCTGTGTAAGCCCTTAAAGGCACTCGTACTAGTTGGATCATGAA 848
DB 463 AGCACTTGTCTTCTAGCTTGGCAAGACCCCAAGGCTCCAAACCAATTTGGATCATGCA 522
QY 849 CGAGTATCGTCTTCGCGAC 867
DB 523 CGAGTATCGCTTCTCGAC 541

RESULT 14
US-10-425-114-29320
; Sequence 29320, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E

QY 593 ATGTTTAAATCATAGTATATGTCAGCTATGCGCGGATAGAGAGAGAGAGTGT 652
 Db 17271 ATAAATATATAAATATATTTATATATCTATATATAATAATAT--ATAAATATATTTAT 17215
 QY 653 ACTTCTATGTGCAAGAGATCGAAATATAGAAATGAGATAGACCGAACCAGGTAACGA 712
 Db 17214 ATATATCTATATATAATAATATAAATATATTTATATATATCTATATAATAATAT 17155
 QY 713 CTTGAGGATATTTGGAAGCCAGCGAGCTGATAGGATGATCAGATCGGAGACTTCTCGGC 772
 Db 17154 ATAAATATATTTATATATCTATATAATAATATAAATATATAAATATATTTATATATATCTATA 17095
 QY 773 CTATCGGATTAAGAAACCCCTAGTTTCTCTACTCTGTTAAAGCCCTAAAGGCACCTCGTA 832
 Db 17094 ATATATAAATATATAAATTTATATATATCTATATATAATAATATAAATATATAATCTTA 17035
 QY 833 CTAGTTGGATCATGAACGAGTATCGTCTTCCGACCATGAACCGGAGAGTACCAAAAGG 892
 Db 17034 TCATATATA-----AATATACAAATATATATCTATCATTA 17001
 QY 893 TATAAATCTCATATAAATCTATATATATCTTATTCATATACATACATAGATATAACCTTAG 952
 Db 17000 TATAAATACAAATATATTTATATATATCTATCATATATAAATATAAATATAAATATTTAT 16941
 QY 953 CTAGGTGGTGGCCCTTTAAATTTGAATTAATCCCTAGACAGTTTGAATTTTCTTTT 1012
 Db 16940 ATATATCTATCATATATAAATATACAAATATATTTATATATCTATCATATATAATA 16881
 QY 1013 TTGACTAGTTTATTTATTTATTTTGGAAATTCGATCGATAAGATCAAAATATCTTGTGAA 1072
 Db 16880 T--AATGTATTTATATAAATATATAATATATATATATATATATATATATATATATAT 16823
 QY 1073 TGGACTAAATGTCAGCGGGTGGTGGCTTTAAATCCAGAAAATGTTTCATCATATGGG 1132
 Db 16822 ATCTATAATATATAAATATATAAATATATAAATATGTAATGTAATATATATATATAT 16763
 QY 1133 TGAACCTTTTAAATGCTAGACATGCCCATATGTTATAGTAGATACATATTAATAGATAG 1192
 Db 16762 AGACATATATAATATATATTTAAATATATATTTATATATATATATATATATATATAT 16708
 QY 1193 ATGCATACATATATATAAACAACAGTATCACACTCGACATTCATATACCTTAAATTC 1252
 Db 16707 TAATATATTTATATAATATTAATATATTTATATATATATATATATATATATATATAT 16648
 QY 1253 TGACAGACATAGTTAGTTTCTTCAAAATTTAAGACATGAATGTTCCCTGCTTCTCTCA 1312
 Db 16647 ATAAATAATATATAAATATATTTATTAATATATATATATAAATAATATAAATATATA 16588
 QY 1313 CATTAATTCATGCTCTCTATTTAAGTTTACCAACATTTTGAATAAATTTGGCATATAT 1372
 Db 16587 AAATATATAATATAAATACATATATATATATATATATATATATATATATATATATAT 16528
 QY 1373 GAATATACCAATATTTATATGCGAACAATTTAAATCTATACGAATGATTAACCGTTTA 1432
 Db 16527 TAATATATATAAATATATAATTTAATATATATATATATATATATATATAAATCT 16468
 QY 1433 TGGAGTAGACGAAAAAATATATATGATACGGAATAATGCAATGGATAGATAAATACAT 1492
 Db 16467 TATATAATATTTATATATATATAA--TATTATAAATATATATATATATATATATAT 16411
 QY 1493 TTTTGGGCTCTTCGACTTATATGTCGACCAATTTGAAACCAATTAATTTATAAATTTT 1552
 Db 16410 TATAAATACAAATATATATATATTAATATATAAATATATATATATATATATATATAT 16351
 QY 1553 CTATGAT 1607
 Db 16350 CATTA 16291
 QY 1608 TGACATAAATATCTATGTTGATGTTATGCAATTCCTGCTACTATAAATAAATTA 1667
 Db 16290 TATTATACAT 16231

QY 1668 CGCAATTAATATATAAAAAATAAATAATAAATCATCTTAATATATATTTGCTATTGTAC 1727
 Db 16230 TATAATATATATATAAATATATATATATATAAATATAAATATAAATATAAATATA 16171
 QY 1728 GTCATATAGTACTCTAAATTTCTTCTAAACGTCGTATCTTTTCTGCTAACTAAC 1787
 Db 16170 AATATATAAGATATATATATAAATAAATAAATATATATATATATATATATATAA 16111
 QY 1788 TTTACATAGTTTGTGCAATCTTCTTTCRAA 1816
 Db 16110 AATATATAAAT 16082

RESULT 7
 US-11-096-568A-33561
 ; Sequence 33561, Application US/11096568A
 ; Publication No. US20060048240A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexandrov, Nickolai et al.
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide
 ; FILE OF INVENTION: Therby
 ; FILE REFERENCE: 2750-1592PUS2
 ; CURRENT APPLICATION NUMBER: US/11/096,568A
 ; CURRENT FILING DATE: 2005-04-01
 ; NUMBER OF SEQ ID NOS: 34471
 ; SEQ ID NO 33561
 ; LENGTH: 1263
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)-(1263)
 ; OTHER INFORMATION: Ceres Seq. ID no. 13603447
 ; US-11-096-568A-33561

Query Match 3.4%; Score 88.8; DB 11; Length 1263;
 Best Local Similarity 62.3%; Pred. No. 3.9e-05;
 Matches 157; Conservative 0; Mismatches 92; Indels 3; Gaps 1;

QY 609 ATATGTATATGTCAGCTATGCGGCGATAGGAGAGAGAGTGGTACTTCTATGTGCCAAG 668
 Db 273 ATGGGTTTTTACCAAGTAGGGGTTTATTCGGTGAAGAAAGATGGTATTTTTCAGCCCGAG 332
 QY 669 AGATCGGAATATAGAAATGAGATAGACCGAAGCGAGTAAACGACTTCAGGATATTTGAA 728
 Db 333 GATAGGAAGTATCCAAACGGTCAAGACCTATCGGGTTCGGGGTGGGTATTTGAA 392
 QY 729 AGCCACCGGAGCTGATAGGATGATCAGATCGGAGACTTCTCGGCCTATCGGATTAAGAA 788
 Db 393 AGCCACCGGTACGGATAAAGTTATCTCGACGAG--GGAAGAAGAGTTGGTATCAAGAA 449
 QY 789 AACCTAGTTTCTACTCTGTTAAAGCCCTAAAGGCACTCGTACTAGTTGGATCATGAA 848
 Db 450 AGCTTTGGTGTTTTACATTTGGAAGAGCTCCAAAGAGAACCAAAACCAATTTGGATTATGCA 509

QY 849 CGAGTATCGTCT 860
 Db 510 TGAGTACCGTCT 521

RESULT 8
 US-11-096-568A-5850
 ; Sequence 5850, Application US/11096568A
 ; Publication No. US20060048240A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexandrov, Nickolai et al.
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide
 ; FILE OF INVENTION: Therby
 ; FILE REFERENCE: 2750-1592PUS2
 ; CURRENT APPLICATION NUMBER: US/11/096,568A
 ; CURRENT FILING DATE: 2005-04-01
 ; NUMBER OF SEQ ID NOS: 34471
 ; SEQ ID NO 5850

LENGTH: 1396

TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(1396)
OTHER INFORMATION: Ceres Seq. ID no. 14312047

US-11-096-568A-5850

Query Match 3.3%; Score 86; DB 11; Length 1396;

Best Local Similarity 63.6%; Pred. No. 0.00011;

Matches 150; Conservative 0; Mismatches 80; Indels 6; Gaps 1;

Qy 637 GGAGAGAAAGTGGTACTTCTATGCGCAAGAGATCGGAATATAGAAATCGAGATAGA 696

Db 338 GCGGAGAAAGTGGTACTTCTTCTCTCGGACCGGAAATACCCGCAAGTTCACGG 397

Qy 697 CCGAACCGAGTAACGACTTCAGGATATTGGAAGCCACCGAGCTGATAGATATAGA 756

Db 398 CCGAACCGCGCGCGGAAAGCGGCTATTGGAAGCCACCGCGGCTGATAAACCAGATCGA 457

Qy 757 TCGGAGACTTCTCGGCTATCGGATTAAGAAACCCTAGTTTCTACTCTGSTARAGCC 816

Db 458 AAACCGAAAGCGC-----TTGGGATCAAGAAAGCTCTGGTTTTCAGCCGGAAGGCC 511

Qy 817 CCTAAGGCACTCGTACTAGTTGGATCATGAACGAGTATCGTCTTCCGCAACCATGA 872

Db 512 CCCAAGGTGTGAACCAATGGATTATGCAAGNATATCGCTCGCCCAATGTTGA 567

RESULT 9

US-11-096-568A-7102

Sequence 7102, Application US/11096568A

Publication No. US20060048240A1

GENERAL INFORMATION:

APPLICANT: Alexandrov, Nikolai et al.

TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides

FILE OF INVENTION: Therby

FILE REFERENCE: 2750-1592PUS2

CURRENT APPLICATION NUMBER: US/11/096.568A

CURRENT FILING DATE: 2005-04-01

NUMBER OF SEQ ID NOS: 34471

SEQ ID NO 7102

LENGTH: 1395

TYPE: DNA

ORGANISM: Glycine max

FEATURE:

NAME/KEY: misc feature

LOCATION: (1)..(1395)

OTHER INFORMATION: Ceres Seq. ID no. 15170403

US-11-096-568A-7102

Query Match 3.2%; Score 83.4; DB 11; Length 1395;

Best Local Similarity 63.0%; Pred. No. 0.00029;

Matches 148; Conservative 0; Mismatches 81; Indels 6; Gaps 1;

Qy 626 TGGCGGCGATAGGAGAGAGTGGTACTTCTATGTCGCAAGAGATCGGAATATAGAA 685

Db 243 TGGTTCTTACGAGAGAGAGTGGTACTTCTTTTACACCGGACCGGAAATACCGGA 302

Qy 686 ATGAGATAGACCGAACCGAGTAAACGACTTCAGGATATTGGAAGCCACCGGAGTGATA 745

Db 303 ACGTTCCAGGCCGAAACCGGCGGCGGGAACCGGTTACTGGAAGCAACCGGGGGGNTA 362

Qy 746 GGATGATCAGATCGGAGACTTCTCGGCTATCGGATTAAGAAACCCCTAGTTTCTACT 805

Db 363 AGCCCAT-----TGGCCACCCCAACCGGTTGGATATAAAAAAGCTTGGTGTATTACG 416

Qy 806 CTGTAAGCCCTTAAGGCACTCGTACTAGTTGGATCATGAACGAGTATCGTCT 860

Db 417 CAGGGAAGCTCCGAAGGGGACAAGCAATGGATCATGACGAGTATCGTCT 471

RESULT 10

US-11-096-568A-4700

Sequence 4700, Application US/11096568A

Publication No. US20060048240A1

GENERAL INFORMATION:

APPLICANT: Alexandrov, Nikolai et al.

TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides

FILE OF INVENTION: Therby

FILE REFERENCE: 2750-1592PUS2

CURRENT APPLICATION NUMBER: US/11/096.568A

CURRENT FILING DATE: 2005-04-01

NUMBER OF SEQ ID NOS: 34471

SEQ ID NO 4700

LENGTH: 1191

TYPE: DNA

ORGANISM: Glycine max

FEATURE:

NAME/KEY: misc feature

LOCATION: (1)..(1191)

OTHER INFORMATION: Ceres Seq. ID no. 14303479

US-11-096-568A-4700

Query Match 3.2%; Score 82.6; DB 11; Length 1191;

Best Local Similarity 60.8%; Pred. No. 0.00038;

Matches 155; Conservative 0; Mismatches 94; Indels 6; Gaps 1;

Qy 618 TGCAGCTATGCGCGCATAGAGAGAGAAAGTGGTACTTCTATGTCGCAAGAGATCGAA 677

Db 246 TCAGGATGCTTTGTACGAGAGAAAGTGGTATTTTTTTCACGCGGAGGACCGCA 305

Qy 678 ATATAGAAATGAGATAGACCGAACCGAGTAACGACTTCAGGATATTGGAAGCCACCGG 737

Db 306 GTACCCGACGGTTCGCGGCGGAAACCGGTCGCGGAAACCGGTTACTGGAAGGCAACCGG 365

Qy 738 AGCTGATAGATCATGATCGGAGACTTTCGCGCTATCGGATTAAGAAACCCCTAGT 797

Db 366 AGCGGAT-----AAACCACTTGGTAAACCGGTTGGGATCAAGAAAGCGTTGGT 419

Qy 798 TTCTACTCTGTTAAAGCCCTTAAAGCACTCGTACTAGTTGGATCATGAACGAGTATCG 857

Db 420 TTTTACCTGGAAGAGCGCCCAAGGAGAGAAACTACTGGATCATGACGAGTATCG 479

Qy 858 TCTTCCGACCATGA 872

Db 480 TCTTGACAGCTGA 494

RESULT 11

US-11-121-086-25

Sequence 25, Application US/11121086

Publication No. US20050266459A1

GENERAL INFORMATION:

APPLICANT: POULSEN, TIM S.

APPLICANT: NIELSEN, KIRSTEN V.

TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES

FILE REFERENCE: 09138.6000-00000

CURRENT APPLICATION NUMBER: US/11/121.086

CURRENT FILING DATE: 2005-05-04

PRIOR APPLICATION NUMBER: 60/567,570

PRIOR FILING DATE: 2004-05-04

NUMBER OF SEQ ID NOS: 107

SOFTWARE: PatentIn version 3.3

SEQ ID NO 25

LENGTH: 173602

TYPE: DNA

ORGANISM: Homo sapiens

US-11-121-086-25

Query Match 3.2%; Score 82.2; DB 14; Length 173602;

Best Local Similarity 48.0%; Pred. No. 0.0012;

Matches 393; Conservative 0; Mismatches 413; Indels 13; Gaps 5;

Qy 1126 ATATGCGTGAACCTTTTAAATTTGTAGACATGCGCCCATATGTTATAGTAACATTA 1185

RESULT 14

US-10-301-480-568794/c
; Sequence 568794, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 568794
; LENGTH: 990
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-568794

Query Match 3.1%; Score 80.6; DB 10; Length 990;
Best Local Similarity 47.4%; Pred. No. 0.00077;
Matches 337; Conservative 0; Mismatches 369; Indels 5; Gaps 3;
QY 1173 TAGAATCAATTAATAGATAGATGCATACACATATATATAAACAACAGTATACACACTCG 1232
DB 752 TATGCTACATTTATATATGGATACATATAAACAATATAATATAAATAAATTT 693
QY 1233 ACATTTCATATACCTTAATCTGCAGACATAGTTAGTTTTCTTACAAATTTATGACATG 1292
DB 692 AAATGACAAATAATGTTATATATATATATATATATATATATATATATATATATAT 633
QY 1293 AATGTTCTGCTCTTCCATCAATTAATTCATGCTCTTCTTAAAGTTACCCAAATTTTT 1352
DB 632 TAATTTATATATAAATAATACATTTTAAATAATATAATATAATATAATATAATAT 573
QY 1353 TGAATAAATTTGGCATATATGAAT-TATACCAACATATTTATATATCGGAACATTTAAATC 1411
DB 572 AAATTTAT 513
QY 1412 TATACGAATGATAACGGTTTATGGAGTAGACCGAAAAAATAATATATATATATATAT 1471
DB 512 TAATACAAAT 453
QY 1472 CAATGGATAGATAAATACATTTTGGGCTCTTTGCGACTTATATGTCGCACCATTTGAA 1531
DB 452 TATTATAAATAATATAATTAATGTTACATTTAAATAATATAATATATATATATATAC 393
QY 1532 ACCATAAATTTATAAATAATTTCTATGATATATATATATATATATATATATATATAT 1590
DB 392 ATTATTTAT 333
QY 1591 AGCTAAACACACGGGTTGACATATATCTATGTTGATGTTATGTCACATTCCTTGAC 1650
DB 332 AATATAAATAAT 273
QY 1651 TAATAAATAAATAATTAACCAATTAATAAATAAATAAATAAATAAATAAATAAATAA 1710
DB 272 ATATAAATAAATAAT 213
QY 1711 TATATTTGCAATGTTACGTCATATGATAG---TACTCTAAATTTCTCTAAACCGTCTAT 1767
DB 212 TATTTTAT 153
QY 1768 CTTTTTTTGGCTTAAGCTTAACCTTTACATAGTTTGTGAACTCTTTTCAAACCATATCTTC 1827
DB 152 ATAG 93
QY 1828 GATAAATGATATTTTTCATAGATATGTTAGTCTATATTTGCAATTTGAT 1878

DB 92

TAGATACAATATATATACATATATAATAACAAGTAAATTTATATATAATAAAT 42

RESULT 15

US-10-301-480-1182203/c
; Sequence 1182203, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1182203
; LENGTH: 990
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1182203

Query Match 3.1%; Score 80.6; DB 10; Length 990;
Best Local Similarity 47.4%; Pred. No. 0.00077;
Matches 337; Conservative 0; Mismatches 369; Indels 5; Gaps 3;
QY 1173 TAGAATCAATTAATAGATAGATGCATACATATATAAACAACAGTATACACACTCG 1232
DB 752 TATGCTACATTTATATATGGATACATATAAACAATATAATATAAATAAATTT 693
QY 1233 ACATTTCATATACCTTAATCTGCAGACATAGTTAGTTTTCTTACAAATTTATGACATG 1292
DB 692 AAATGACAAATAATGTTATATATATATATATATATATATATATATATATATATAT 633
QY 1293 AATGTTCTGCTCTTCCATCAATTAATTCATGCTCTTCTTAAAGTTACCCAAATTTTT 1352
DB 632 TAATTTATATATAAATAATACATTTTAAATAATATAATATAATATAATATAATAT 573
QY 1353 TGAATAAATTTGGCATATATGAAT-TATACCAACATATTTATATCGGAACATTTAAATC 1411
DB 572 AAATTTAT 513
QY 1412 TATACGAATGATAACGGTTTATGGAGTAGACCGAAAAAATAATATATATATATATAT 1471
DB 512 TAATACAAAT 453
QY 1472 CAATGGATAGATAAATACATTTTGGGCTCTTTGCGACTTATATGTCGCACCATTTGAA 1531
DB 452 TATTATAAATAATATAATTAATGTTACATTTAAATAATATAATATATATATATATAC 393
QY 1532 ACCATAAATTTATAAATAATTTCTATGATATATATATATATATATATATATATATAT 1590
DB 392 ATTATTTAT 333
QY 1591 AGCTAAACACACGGGTTGACATATATCTATGTTGATGTTATGTCACATTCCTTGAC 1650
DB 332 AATATAAATAAT 273
QY 1651 TAATAAATAAATAATTAACCAATTAATAAATAAATAAATAAATAAATAAATAAATAA 1710
DB 272 ATATAAATAAATAAT 213
QY 1711 TATATTTGCAATGTTACGTCATATGATAG---TACTCTAAATTTCTCTAAACCGTCTAT 1767
DB 212 TATTTTAT 153
QY 1768 CTTTTTTTGGCTTAAGCTTAACCTTTACATAGTTTGTGAACTCTTTTCAAACCATATCTTC 1827
DB 152 ATAG 93

Qy 1828 GATAAATGATATTTTCATAGATATTCCTAGTCTCTATATTTGATAATTGAT 1878
Db 92 TAGATACAATATTTACATATATAATAACAAGTAAATTTATATATAATAAAT 42

Search completed: April 6, 2006, 09:58:15
Job time : 1165 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 6, 2006, 10:24:48 ; Search time 580 Seconds
(without alignments)

2614.156 Million cell updates/sec

Title: US-10-780-703-2

Perfect score: 2034

Sequence: 1 MAVSSTSIIPMSNQVNNN.....MWNPIVDGNGRDHYTNIPFK 379

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9267905 seqs, 2000278028 residues

Total number of hits satisfying chosen parameters: 18535810

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh

-Q=/abses/ABSSWEB.spool/US10780703/runat_06042006_101351_6632/app_query.fasta_1

-DB=Published Applications NA New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1

-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62

-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100

-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0

-MAXLEN=200000000 -HOST=abses05h

-USER=US10780703 @CGN_1_1_652 -runat_06042006_101351_6632 -NCPU=6 -ICPU=3

-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMSOUP=120

-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7

-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:

1: /SID55/ptodata/2/pubpna/US08 NEW PUB.seq.*
2: /SID55/ptodata/2/pubpna/US06 NEW PUB.seq.*
3: /SID55/ptodata/2/pubpna/US07 NEW PUB.seq.*
4: /SID55/ptodata/2/pubpna/PCT NEW PUB.seq.*
5: /SID55/ptodata/2/pubpna/US03 NEW PUB.seq.*
6: /SID55/ptodata/2/pubpna/US09 NEW PUB.seq.*
7: /SID55/ptodata/2/pubpna/US10 NEW PUB.seq.*
8: /SID55/ptodata/2/pubpna/US10 NEW PUB.seq.*
9: /SID55/ptodata/2/pubpna/US10 NEW PUB.seq.*
10: /SID55/ptodata/2/pubpna/US10 NEW PUB.seq.*
11: /SID55/ptodata/2/pubpna/US11 NEW PUB.seq.*
12: /SID55/ptodata/2/pubpna/US11 NEW PUB.seq.*
13: /SID55/ptodata/2/pubpna/US11 NEW PUB.seq.*
14: /SID55/ptodata/2/pubpna/US11 NEW PUB.seq.*
15: /SID55/ptodata/2/pubpna/US60 NEW PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--|
| 1 | 503 | 24.7 | 1263 | 11 | US-11-096-568A-33561 Sequence 33561, A |
| 2 | 497 | 24.4 | 1606 | 11 | US-11-172-740-1242 Sequence 1242, Ap |

| | 3 | 494 | 24.3 | 1219 | 11 | US-11-096-568A-31996 |
|----------|-------|------|------|------|----------------------|----------------------|
| Sequence | 488.5 | 24.0 | 1191 | 11 | US-11-096-568A-4700 | |
| Sequence | 482.5 | 23.7 | 963 | 11 | US-11-087-099-65 | |
| Sequence | 479.5 | 23.6 | 1304 | 11 | US-11-096-568A-7863 | |
| Sequence | 478 | 23.5 | 1396 | 11 | US-11-096-568A-5850 | |
| Sequence | 474.5 | 23.3 | 1423 | 8 | US-10-391-414-4 | |
| Sequence | 473.5 | 23.3 | 1395 | 11 | US-11-096-568A-7102 | |
| Sequence | 471 | 23.2 | 1189 | 11 | US-11-096-568A-26116 | |
| Sequence | 463 | 22.8 | 1409 | 11 | US-11-096-568A-19208 | |
| Sequence | 453.5 | 22.3 | 1218 | 11 | US-11-087-099-50 | |
| Sequence | 448.5 | 22.1 | 1833 | 11 | US-11-096-568A-24275 | |
| Sequence | 447.5 | 22.0 | 1393 | 11 | US-11-096-568A-22731 | |
| Sequence | 444.5 | 21.9 | 897 | 11 | US-11-096-568A-30488 | |
| Sequence | 441.5 | 21.7 | 1316 | 11 | US-11-096-568A-23342 | |
| Sequence | 434 | 21.3 | 852 | 11 | US-11-096-568A-28895 | |
| Sequence | 432 | 21.2 | 493 | 11 | US-11-116-881A-790 | |
| Sequence | 430.5 | 21.2 | 1388 | 11 | US-11-096-568A-20504 | |
| Sequence | 427.5 | 21.0 | 1412 | 11 | US-11-096-568A-3792 | |
| Sequence | 424.5 | 20.9 | 1478 | 11 | US-11-096-568A-19423 | |
| Sequence | 423 | 20.8 | 1787 | 11 | US-11-096-568A-3133 | |
| Sequence | 421 | 20.7 | 2372 | 11 | US-11-096-568A-24718 | |
| Sequence | 419.5 | 20.6 | 1192 | 11 | US-11-096-568A-20705 | |
| Sequence | 411 | 20.2 | 1298 | 11 | US-11-096-568A-6578 | |
| Sequence | 410.5 | 20.2 | 2122 | 11 | US-11-096-568A-31731 | |
| Sequence | 403.5 | 19.8 | 1528 | 11 | US-11-096-568A-17746 | |
| Sequence | 403 | 19.8 | 1374 | 11 | US-11-096-568A-13595 | |
| Sequence | 400 | 19.7 | 918 | 11 | US-11-096-568A-29126 | |
| Sequence | 399.5 | 19.6 | 1203 | 11 | US-11-096-568A-8631 | |
| Sequence | 398 | 19.6 | 1419 | 11 | US-11-096-568A-31178 | |
| Sequence | 397.5 | 19.5 | 1251 | 11 | US-11-096-568A-17565 | |
| Sequence | 394 | 19.4 | 1648 | 11 | US-11-096-568A-25426 | |
| Sequence | 391.5 | 19.2 | 879 | 11 | US-11-096-568A-30325 | |
| Sequence | 388 | 19.1 | 596 | 11 | US-11-096-568A-26867 | |
| Sequence | 385 | 18.9 | 1227 | 11 | US-11-096-568A-34447 | |
| Sequence | 380 | 18.7 | 1353 | 11 | US-11-096-568A-21255 | |
| Sequence | 377.5 | 18.6 | 1370 | 11 | US-11-096-568A-17121 | |
| Sequence | 372.5 | 18.3 | 1285 | 9 | US-10-714-887-261 | |
| Sequence | 372.5 | 18.3 | 1387 | 11 | US-11-096-568A-33784 | |
| Sequence | 372.5 | 18.3 | 1489 | 11 | US-11-096-568A-22192 | |
| Sequence | 358.5 | 17.6 | 1421 | 11 | US-11-096-568A-22904 | |
| Sequence | 350.5 | 17.2 | 1421 | 9 | US-10-714-887-241 | |
| Sequence | 346 | 17.0 | 1371 | 1008 | 11 | |
| Sequence | 345 | 17.0 | 1008 | 11 | US-11-096-568A-28163 | |
| Sequence | 344.5 | 16.9 | 1008 | 9 | US-10-714-887-269 | |

ALIGNMENTS

RESULT 1
US-11-096-568A-33561
; Sequence 33561, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide
; FILE OF INVENTION: Thery
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096, 568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 33561
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1263)
; OTHER INFORMATION: Ceres Seq. ID no. 13603447
US-11-096-568A-33561

Alignment Scores:
Pred. No.: 3,09e-42 Length: 1263
Score: 503.00 Matches: 115
Percent Similarity: 54.8% Conservative: 40

Best Local Similarity: 40.6% Mismatches: 74
 Query Match: 24.7% Indels: 54
 DB: 11 Gaps: 9

US-10-780-703-2 (1-379) x US-11-096-568A-33561 (1-1263)

QY 53 ProGlyPheArgPheHisProThrGluGluGluLeuLeuGluPhePhePheArgArgLys 72
 DB 151 CCGGTTTCGGTTTATCCGACTGACGAGAGCTGATGGTGAATATCTCTGTAGAAAA 210
 QY 73 ValGluGlyLysArgPheAsnValGluLeuLeuThrPheLeuAspLeuThrArgPhe 92
 DB 211 GCCGCGGTACGACTCTCTCTCCAGCTCATAGCTGAATCGATCTCTACAACTTGTAT 270
 QY 93 ProThrGluLeuProAlaMetAlaAlaLeuGlyGlyLysGluThrPhePheValPro 112
 DB 271 CCATCGGTTTACCAAGTAAGCGGTATTTCGGTGAAGAAATGGTATTTTTCAGCCCG 330
 QY 113 ArgAspArgLysPheArgAsnGlyAspArgProAsnArgValThrSerGlyThrTrp 132
 DB 331 AGGATAGAGAGTATCCAAACGGTCAAGACCTATCGGTTCCGGGTCGGGTTATGG 390
 QY 133 LysAlaThrGlyAlaAspArgMetLeuArgSerGluThrSerArgProLeuGlyLys 152
 DB 391 AAAGCCACCGGTACGATAAGTATCTCGACGGAG---GGAAGAAAGAGTTGGTATCAAG 447
 QY 153 LysThrLeuValPhePheSerGlyLysAlaProLysGlyThrArgThrSerTrpLeuMet 172
 DB 448 AAAGCTTTGGTGTATACATTCGAAAGCTCCAAAGGAACCAACCAATTTGGATTATG 507
 QY 173 AsnGluThrArgLeu-----ProHisHisGluThrGluLysGlnLysAlaGluLeu 190
 DB 508 CATGAGTACCGTCTCATCGAACCTCTCGTGAAGTGAAGCAGCAAGCTTGTATGAG 567
 QY 191 SerLeuCysArgValThrLysArgProGlyValGluAspHisProSerValProArgSer 210
 DB 568 GTTTTATGTCGAATATACAAAAGCAACAAAGCGCACAAACAAAGCTTACAAATCTA 627
 QY 211 LeuSerThr-----ArgHisHisAsnHisAsnSerSerThrSerSer----- 224
 DB 628 ATGACGAGTGGTGTGTAATACACAAATATGTTCTGTCGACATCTTCTTCGTCATCAA 687
 QY 225 -----ArgLeuAlaLeuArgGlnGlnHisHisSer 235
 DB 688 TAGCAGCAGCTCTCGAGTGGTGTGATGATGATGACAAAGAGTTGGGTTTCCCGCC 747
 QY 236 SerSerSerAsn-----HisSerAspAsnAsnLeuAsnAsnAsnAsnAsnAsn 251
 DB 748 GTTTCATCAACCGCGTGCCTCATAGTCATAGCCGTTTAAACCAATCATATAAACCGGG 807
 QY 252 -----AsnAsn 253
 DB 808 TTTCCAGGTTTAGCCAGGAGCCAGTTTGTATGGCGGAATTTGATTGGACAGAACTCG 867
 QY 254 LeuGluLysLeuSerThrGluThrSer-----GlyAspGlySerThr 267
 DB 868 GTCCCGAACTCGGAGTACAGTACAGTTCGAGTATTCGTTACGCGTACCGTGAAGC 927
 QY 268 ThrThrThrThr-----AsnSerAsnSerAspValThrThrAlaLeu 282
 DB 928 CAGCAACAACCTCAGGGGATTCCTCGGTTTAAATAATACTCGGACCTCTCG----- 978
 QY 283 AlaAsnGln 285
 DB 979 GCTAATCAG 987

RESULT 2

US-11-172-740-1242
 ; Sequence 1242, Application US/11172740
 ; Publication No. US20060057724A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MASCIA, Peter
 ; APPLICANT: ALEXANDROV, Nikolai

APPLICANT: BROVER, Vyacheslav
 TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND POLYPEPTIDES ENCODED THEREBY USEFUL FO
 FILE OF INVENTION: PLANT CHARACTERISTICS AND PHENOTYPES
 FILE REFERENCE: 2750-1602PUS2
 CURRENT APPLICATION NUMBER: US/11/172,740
 CURRENT FILING DATE: 2005-06-30
 PRIOR APPLICATION NUMBER: 60/583,621
 PRIOR FILING DATE: 2004-06-30
 PRIOR APPLICATION NUMBER: 60/584,829
 PRIOR FILING DATE: 2004-06-30
 PRIOR APPLICATION NUMBER: 60/584,800
 PRIOR FILING DATE: 2004-06-30
 NUMBER OF SEQ ID NOS: 2523
 SEQ ID NO 1242
 LENGTH: 1606
 TYPE: DNA
 ORGANISM: Arabidopsis thaliana
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (1)..(1606)
 OTHER INFORMATION: Ceres CLONE ID no. 38214
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (1)..(1606)
 OTHER INFORMATION: Also known as Ceres CDNA ID no. 23373803
 US-11-172-740-1242

Alignment Scores:
 Pred. No.: 1,74e-41 Length: 1606
 Score: 497.00 Matches: 124
 Percent Similarity: 45.0% Conservative: 47
 Best Local Similarity: 32.6% Mismatches: 109
 Query Match: 24.4% Indels: 100
 DB: 11 Gaps: 12

US-10-780-703-2 (1-379) x US-11-172-740-1242 (1-1606)

QY 24 IleGluAspAsnAspHisArgGlyGlyGlnGluSerHisValGlnAsnGluAspGluAla 43
 DB 147 ATGAAAGACCGATCTTCGGTGGTCCACCACCGCCACCACTAAC----- 194
 QY 44 AspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThrGluGluGlu 63
 DB 195 -----CTTCTCCAGCTTCGGTTTCCCTCACCCTACCACGAAAGAG 233
 QY 64 LeuLeuGluPhePheLeuArgArgLysValGluGlyLysArgPheAsnValGluLeuLeu 83
 DB 234 CTTGTGTTCATCTACCTCAACGCGAAAGCAGCCTCTGCTCTTACCTGTCCGCATATC 293
 QY 84 ThrPheLeuAspLeuThrArgPhePheProThrGluLeuProAlaMetAlaAlaLeuGly 103
 DB 294 GCCGAGTCGATCTCTATTAATTTGATCCATGGGAATCTCCGCTAAAGCATCGTTTGA 353
 QY 104 GluLysGluThrPhePhePheValProArgAspArgLysThrArgAsnGlyAspArgPro 123
 DB 354 GAACAAGAATGGTACTCTTTAGTCCACGAGATCGAAGATATCCAAACGAGGACCAAGCA 413
 QY 124 AsnArgValThrThrSerGlyThrTrpLysAlaThrGlyAlaAspArg---MetileArg 142
 DB 414 AACAGAGCGGCGACTTCAGGTTATTGGAAGCGCCGCTACAGATAAACCCGCTACTTGT 473
 QY 143 SerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheThrSerGlyLysAla 162
 DB 474 TCCGACCGGTAACCAAAAGTGGCGGTGAAGAGGCGACTAGTCTTCTACAGTGGTAACCA 533
 QY 163 ProLysGlyThrArgThrSerTrpIleMetAsnGluThrArgLeu-----Pro 178
 DB 534 CCAAAAGCGGTTAAAGTATGGATCATGATGATATCGTCTCATCGAAACCAACCA 593
 QY 179 His-----HisGluThrGluLysThrGlnLysAlaGlu 189
 DB 594 AACATCGACCTCTGCTGTGATTTTCGGCAACAAAGAAAACTCCTCAGACTTGATGAT 653

US-10-780-703-2 (1-379) x US-11-096-568A-4700 (1-1191)

```
QY 53 ProGlyPheArgPheHisProThrGluGluGluLeuLeuGluPheTyrLeuArgArgLys 72
DB 115 CTGGATTTCAGATTCATCAACAGATGAAGAACTCGTCATCACTACCTTTGTGCGCAA 174
QY 73 ValGluGlyLeuArgPheAsnValGluLeuLeuThrPheLeuAspLeuTyrArgTyrAsp 92
DB 175 TGGCGTTCCGACATATCGCGTTCCCAATAATCCGGAATGATCTGTACAAGTACGAC 234
QY 93 ProTrpGluLeuProAlaMetAlaAlaLeuGlyGluLeuTrpTyrPheTyrValPro 112
DB 235 CCTTGGACCTTCCAGGAATGGCTTTGTACGAGAGAAAGTGGTATTTTTCACGCCG 294
QY 113 ArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrp 132
DB 295 AGGACCGCAAGTACCGAACGTTCCGCGCCGACCGGTCCCGGGAACCGGGTACTGG 354
QY 133 LysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLys 152
DB 355 AAGGCAACCGGAGCGGATAACACGATGGTAA-----CCGAAACCGTTGGGATCAAG 408
QY 153 LysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMet 172
DB 409 AAAGCGTTGGTTTTTACGCTGGAAAGCGCCCAAGGAGAGAAACTAACTGGATCATG 468
QY 173 AsnGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLysAlaGluLeuSer--- 191
DB 469 CACGAGTATCGTCTTCAGACGCTGGATCGTTCCGTTCCGCAAAAGAACAGCTTAAGCGTG 528
QY 192 -----LeuCybArgValTyrLysArgProGly----- 200
DB 529 GATGACTGGTCTGTGTCGCAATTTACACAGAAAGGTGCAATTCGAAAGCAACACTA 588
QY 201 -----ValGluAspHis----- 204
DB 589 CCACCAACCGAGTGGGTCCGCAAAATTGAATGTTCCGAAATGCGAGCAGAGCGGAG 648
QY 205 -----ProSerValProArgSerLeuSerThr----- 213
DB 649 ATTCTGCCCGCAGATCCGCGGTATACGCGCGACGCGTGGCGGATTCCTGACTTCGAG 708
QY 214 -----ArgHisHisAsnHisAsnSerSerThrSerArgLeuAla 227
DB 709 GCTTCGCACTCGGTGCGCGCGTGCACACGACGAGTTCGAGCTGCCGACGAGTGGTG 768
QY 228 LeuArgGlnGlnHisHisSerSerSerSerAsnHisSerAspAsnAsnLeuAsnAsn 247
DB 769 TCGGCGGAGTTTCGAGCGAGGTGCAGAGCGAGCCGAGAGGGGGCAGCAACAACAAC 828
QY 248 -----AsnAsnAsnIleAsnAsnLeuGlu 255
DB 829 GAGTTTGCATATAATTACGTGGATGCCACTCTCCGGAATAATCAGATGTCGCGCTGCAG 888
QY 256 Lys-----LeuSerThrGluTyrSerGlyAspGlySerThrThrThrThrThr 271
DB 889 GATATTTTCATGTACCTCTCCAAAGTCTCTCGCAATTA-ACAGCACCACCCACGACG 947
QY 272 ThrAsnSerAsnSerAspValThrIleAlaLeuAlaAsnGlnAsnIleTyrArgProMet 291
DB 948 AGACGAGCTGCTGTGTGTGTGTATGCATT-----ATG 983
QY 292 ProTyrAspThrSerAsnAsnThrLeuIleValSerThrArgAsnHisGlnAspAsp 311
DB 984 CATTTTGGCAGTAGCTCTCACTGCACCACTTTGCA---CCGAAGCAGCAGACTGAGTGA 1040
QY 312 GluThrAlaIleValAspAspLeuGlnArgLeuValAsnTyrGlnIleSerAspGlyAla 331
DB 1041 AGATCCGTGAGTGTGTCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1100
QY 332 ThrThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsnMetIle 347
DB 1101 AGCACCAGATGAAGAGAGAAAGATCTCTAT---ATGACATGATG 1145
```

RESULT 5

```
US-11-087-099-65
; Sequence 85, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE OF INVENTION: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 65
; LENGTH: 963
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-11-087-099-65
```

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 3,066-40 | Length: | 963 |
| Score: | 482.50 | Matches: | 100 |
| Percent Similarity: | 60.0% | Conservative: | 29 |
| Best Local Similarity: | 46.5% | Mismatches: | 53 |
| Query Match: | 23.7% | Indels: | 34 |
| DB: | 11 | Gaps: | 5 |

US-10-780-703-2 (1-379) x US-11-087-099-65 (1-963)

```
QY 24 IleGluAspAsnAspHisArgGlyGlnGluSerHisValGlnAsnGluaspGluAla 43
DB 1 ATGGAGAGTACAGATTCTTCGGTGGTCTCTCCGCGCGCCGCAACCAAAAC----- 48
QY 44 AspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThrGluGluGlu 63
DB 49 -----CTCCCTCAGAGTTCGGGTTTCATCCCAACAGCAGAGAA 87
QY 64 LeuIleGluPheTyrLeuArgArgLysValGluGlyLysArgPheAsnValGluLeuIle 83
DB 88 CTTGTAATTCATTACCTCAACGCAAGACAGATTCTGTTCTCTTACCAGTCGCGATCATC 147
QY 84 ThrPheLeuAspLeuTyrArgTyrAspProTrpGluLeuProAlaMetAlaIleGly 103
DB 148 GCGCAGCTGTGATCTTTTACAAATTTGATCCATCGGAACCTTCCGCGCAAGCTTCGTTTGA 207
QY 104 GluLysGluTyrTyrPheTyrValProArgAspArgLysTyrArgAsnGlyAspArgPro 123
DB 208 GAACAGAGATGATGATTTTTCAGTCCAGAGATCGAATATATCCCAACGAGCTAGACCT 267
QY 124 AsnArgValThrThrSerGlyTyrTrpLysAlaThrGlyAlaAspArgMetIleArgSer 143
DB 268 AACCGAGCTGCGACTTCGGTTATTGGAAGCGACTGGTACAGATAAACCGGTGATTCA 327
QY 144 Glu-----ThrSerArgProIleGlyLeuLysLysThrLeuValPheTyrSerGly 160
DB 328 ACCGCGGTGGTGGTAGTAGTAAAGAGTGGAGTTAAAAAGGCTCTAGTGTGTTCACAGTGT 387
QY 161 LysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeu----- 177
DB 388 AAACCCACCAAGAGGAGTTAAATCAGATTGATATGCAATATATCGGTAACTGATTAAT 447
QY 178 ----ProHisHis-----GluThrGluLysTyrGlnLysAlaGluIle 190
DB 448 AAACCTACTCATTTTGTGACTTCGCAACAAGAAAACTCTCTCAGGCTTGATGATGG 507
QY 191 SerLeuCybArgValTyrLysArgProGlyValGluAspHisProSerValProArgSer 210
DB 508 GTGTTGTGCTGATCTACAAGAAA-----AACAAATAGT 540
QY 211 LeuSerThrArgHisHisAsnHisAsnSerSerThrSerSerArg 225
DB 541 ACAGCATCTAGACATCATCATCATCTT-TCATCATATTTCATCTAGA 584
```

RESULT 6

US-11-096-568A-7863
; Sequence 7863, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 7863
; LENGTH: 1304
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1304)
; OTHER INFORMATION: Ceres Seq. ID no. 15176333
US-11-096-568A-7863
Alignment Scores:
Pred. No.: 9,05e-40 Length: 1304
Score: 479.50 Matches: 99
Percent Similarity: 59.1% Conservative: 31
Best Local Similarity: 45.0% Mismatches: 73
Query Match: 23.6% Indels: 17
Gaps: 5
DB: 11

US-10-780-703-2 (1-379) x US-11-096-568A-7863 (1-1304)
QY 18 AsnAsnAsnGluLysGlyIleGluAspAsnAspHisArgGlyGlyGlnGluSerHisVal 37
DB 132 ATTCGACGAGAGAGATCGAGAGATGACGCGCGCGCGCGGAGGATCG----- 185
QY 38 GlnAsnGluAspGluAlaAspAspHisAspHisMetValMet---ProGlyPheArg 56
DB 186 -----GCGGCGACGCGGCGGCGAGGATCTGCAGCTCGCGCGGGGTTTCAGA 233
QY 57 PheHisProThrGluGluGluLeuLeuGluPheThrLeuArgArgLysValGluGlyLys 76
DB 234 TTCATCCGACGCGGAGGAGGAGTGTGTATGACCTACCTGTCGCGCGCGGCGCGGCTG 293
QY 77 ArgPheAsnValGluLeuLeuThrPheLeuAspLeuTyrArgTyrAspProTTPGluLeu 96
DB 294 CCATCTCCGTGCCATCATCGCGGAGGTGACCTCTACAGTTTCGACCCATGCGACCTC 353
QY 97 ProAlaMetAlaAlaIleGlyGlyLysGluTyrPheTyrValProArgAspArgLys 116
DB 354 CCAGATGCGGCTCTACGCGGAGAGAGTGTGTACTTCTTCTCCCGCGCGGCGGCAAG 413
QY 117 TyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTTPLysAlaThrGly 136
DB 414 TACCGAACGGTTCGCGCGCGAACCGGTTCGCGCGGCGGCGGCTACTCGAAGGCGCCGCG 473
QY 137 AlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuVal 156
DB 474 GCGACAAAGCGGCGG-----GGCACCGCCCAAGCCCTTGGCCATCAAGAAGGCGCTCGTC 527
QY 157 PheTyrSerGlyLysAlaProLysGlyThrArgThrSerTTPLeuMetAsnGluTyrArg 176
DB 528 TTCTATGCGGGAGGAGCGCCAGGCGGACAGACCACTGGATCATGACGAGTACCGC 587
QY 177 LeuProHisHisGluThrGluLysTyrGlnLysAlaGluLeuSer----- 191
DB 588 CTCGCGGACGTCGACCGCTCGCCCGCAAGAAAGAACAGCCGCTCAGTTGGATGACTGGGTG 647
QY 192 LeuCysArgValTyrLysArgProGlyValGluAspHisProSerVal-----Pro 208
DB 648 CTTTCCCGGATCTACAAACAAGAGGCGCCCTCGGAGAGGCGCGCGCGGTGACCGGACC 707
QY 209 ArgSerLeuSerThrArgHisHisAsnHisAsnSerSerThrSerSerArgLeuAlaLeu 228
DB 228

DB 708 GCGAGCGCGAGCCCGCGGACGCGCGCTGGCTCGCGCGGAGAGAAACGCGCCCTG 767
RESULT 7
US-11-096-568A-5850
; Sequence 5850, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 5850
; LENGTH: 1396
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1396)
; OTHER INFORMATION: Ceres Seq. ID no. 14312047
US-11-096-568A-5850
Alignment Scores:
Pred. No.: 1,41e-39 Length: 1396
Score: 478.00 Matches: 116
Percent Similarity: 50.1% Conservative: 54
Best Local Similarity: 34.2% Mismatches: 108
Query Match: 23.5% Indels: 61
Gaps: 10
DB: 11

US-10-780-703-2 (1-379) x US-11-096-568A-5850 (1-1396)
QY 18 AsnAsnAsnGluLysGlyIle-----GluAspAsnAspHisArgGlyGlyGlnGlu 34
DB 70 AACACAAACAACAAGCTTGCAAAATTTAGAAATCTTCAGTTTAAAGGAGAAAGAAC 129
QY 35 SerHisValGlnAsnGluAspGluAlaAspAspHisAsp-HisAspMetValMet---Pr 53
DB 130 ACACAAGAAGAAGAAGAAGAACGCGCGCAATGAAGGAGAAATAGAGTTGCCACC 189
QY 53 QGlyPheArgPheHisProThrGluGluGluLeuLeuGluPheThrLeuArgArgLysVa 73
DB 190 TGGTTTCAGATTTCACCCACTGATGAGAATTTGGTGAATCACTACTTGTGTAGGAAGTG 249
QY 73 lGluGlyLysArgPheAsnValGluLeuLeuThrPheLeuAspLeuTyrArgTyrAspPr 93
DB 250 CGCTGTCACCAATCGCGGTTCCGTCATCAAGAGAGTCGATTGTGTACAAAGTTTGATCC 309
QY 93 OTTPGluLeuProAlaMetAlaAlaIleGlyGlyLysGluTyrPheTyrValProAr 113
DB 310 ATGGCAGCTTCCAGAAATTTGGTTTACGCGGAGAAAGATGGTACTTCTTCTCTCTCG 369
QY 113 GAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTTPLY 133
DB 370 GGAACCGGAATACCCGAAACGGTTTCACGCGGACCGCGCGCGGAGCGCTATTGGAA 429
QY 133 sAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLysLY 153
DB 430 AGCCACCGCGCTGTATAAACCCGATCGAAAA-----CCGAAAGCGCTTGGGATCAAGAA 483
QY 153 sThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTTPLeuMetAs 173
DB 484 AGCTCTGGTTTTCGCGGAAAGCCCAAGGTGTGAACCAATTTGGATTATGCA 543
QY 173 nGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLysAlaGluLeuSer----- 191
DB 544 CGAATATCGCTCCGCAATGTTGACCGATCTGCCTCCAGAAAGAAACAAACACTTGAG 603
QY 192 -----LeuCysArgValTyrLysArgProGly---ValGluAspHis-- 204
DB 604 GCTTGATGATTGGGTGTGTGTGATCTACAAACAAGAAAGGAGAGATTGAGAAATACAA 663

QY 205 -----ProSerValProArgSerLeuSerThrArgHisHisAsnHisAs 219
 Db 664 CACAGGCGCAGGAGAGTGAATGTTGAGATGTTTTCATAGTTTGACGACGAGAACGACAG 723
 QY 219 nSerSerThrSerSerArgLeuAlaLeuArgGlnGlnHisHisSerSerSerAs 239
 Db 724 GAAGCCAGAGATTCATAAGCTAGGAATGAGCAATTGTACATGGAGACTTCGGATTGC- 781
 QY 239 nHisSerAspAsnAsnLeuAsnAsnAsnAsnAsnLeuLeuLeuLeuSerTh 259
 Db 782 -----GTGCCAAGGTTGAACAC 798
 QY 259 rGluTyrSerGlyAspGlySerThrThrThrThrThrThrAsnSer----- 274
 Db 799 GGACTCGAGACTCGGACGACGCTGGTTCGCCCGATGTCACGTCGGAGAGGAGGTGCA 858
 QY 275 -----AsnSerAspValThrIleAlaLeuAlaAsnGlnAsnIleTyrAr 289
 Db 859 GAGCGACCCCAAGTCGACGATCTGACCTAAGCTAGAAAC----- 904
 QY 289 gProMetProTyrAspThrSerAsnAsnThrLeuIleValSerThrArgAsnHisGlnAs 309
 Db 905 -----GCGTTTGATTTTACGTTTAATTAATTG-----GA 933
 QY 309 pAspAspGluThrAlaIleValAspAspLeuGlnArgLeuValAsnTyrGlnIle 327
 Db 934 CGATAATAACCTTTCCGTGGATGATACCTTTTGGCACTGTTACAGTATCAATG 988

RESULT 8

US-10-391-414-4
 ; Sequence 4, Application US/10391414
 ; Publication No. US20050278799A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SHINOZAKI, Kazuko
 ; APPLICANT: KATSURA, Koji
 ; APPLICANT: ITO, Yusuke
 ; TITLE OF INVENTION: Stress Induced Promoter Derived From Rice
 ; FILE REFERENCE: 382.1041
 ; CURRENT APPLICATION NUMBER: US/10/391,414
 ; PRIOR FILING DATE: 2003-03-18
 ; PRIOR APPLICATION NUMBER: JP 2002-377316
 ; PRIOR FILING DATE: 2002-12-26
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: PatentIn ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 1423
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 US-10-391-414-4

Alignment Scores:
 Pred. No.: 3,34e-39 Length: 1423
 Score: 474.50 Matches: 93
 Percent Similarity: 62.7% Conservativity: 28
 Best Local Similarity: 48.2% Mismatches: 43
 Query Match: 23.3% Indels: 29
 DB: 8 Gaps: 5

US-10-780-703-2 (1-379) x US-10-391-414-4 (1-1423)

QY 27 AsnAspHisArg-----GlyGlyGlnGluSerHisValGlnAsn 39
 Db 85 GACGATCACCGGAGAGGAGAGAGATGAGCGCGGTCAGGACTCGAGCTGCCG--- 141
 QY 40 GluAspGluAlaAspAspHisAspHisAspMetValMetProGlyPheArgPheHisPro 59
 Db 142 -----CCGGGGTTCGGTTCACCCG 162
 QY 60 ThrGluGluGluLeuIleGluPheTyrLeuArgArgLysValGluGlyLysArgPheAsn 79
 Db 163 ACGGACGAGGAGCTGTGTGATGCACTACTCTCCCGCGGTGCGCGCTCCCATCGCC 222

QY 80 ValGluLeuIleThrPheLeuAspLeuTyrArgTyrAspProTyrGluLeuProAlaMet 99
 Db 223 GTCCCATCATCGCGAGATCGACCTCTACAGTTTCGATCCATGCAGCTTCCCGGATG 282
 QY 100 AlaAlaIleGlyGluLysGluTyrPheTyrValProArgAspArgLysTyrArgAsn 119
 Db 283 CGCGTGTACGAGAGAGAGGTGTTCTCTCTCCCGCGAGACCGCAAGTACCCGAC 342
 QY 120 GlyAspArgProAsnArgValThrThrSerGlyTyrTyrLysAlaThrGlyAlaAspArg 139
 Db 343 GGGTCCGCGCGCAACCCCGCGCGGTTCGGGTACTTGGAAAGCGACCGCGCCGACAG 402
 QY 140 MetIleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTyrSer 159
 Db 403 CCGTGGGCTCG-----CCGAGCCGTTGGCGATCAAGAGGCCCTCGTCTTCTACGCC 456
 QY 160 GlyLysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeuProHis 179
 Db 457 GCGAAGGCGCCCAAGGCGAGAACCACTGGAGTCATGCAAGAGTACCGGCTCGCCGAC 516
 QY 180 HisGluThrGluLysTyrGlnLysAlaGluIleSer-----LeuCysArg 194
 Db 517 GTCACCGCTCCGCCCGCAAGAGAACAGCCTCAGTTGGATGATGGTCTGTGCCGG 576
 QY 195 ValTyr---LysArgProGlyValGluAspHisProSer 206
 Db -577 ATTACAAAGAGGGCGGCTGAGAGAGCGCGCGGCC 615

RESULT 9

US-11-096-568A-7102
 ; Sequence 7102, Application US/11096568A
 ; Publication No. US20060048240A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexandrov, Nikolai et al.
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide
 ; FILE REFERENCE: 2750-1592PUS2
 ; CURRENT APPLICATION NUMBER: US/11/096,568A
 ; CURRENT FILING DATE: 2005-04-01
 ; NUMBER OF SEQ ID NOS: 34471
 ; SEQ ID NO 7102
 ; LENGTH: 1395
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)..(1395)
 ; OTHER INFORMATION: Ceres Seq. ID no. 15170403
 US-11-096-568A-7102

Alignment Scores:
 Pred. No.: 4.14e-39 Length: 1395
 Score: 473.50 Matches: 115
 Percent Similarity: 46.2% Conservativity: 44
 Best Local Similarity: 33.4% Mismatches: 79
 Query Match: 23.3% Indels: 107
 DB: 11 Gaps: 9

US-10-780-703-2 (1-379) x US-11-096-568A-7102 (1-1395)

QY 53 ProGlyPheArgPheHisProThrGluGluGluLeuIleGluPheTyrLeuArgArgLys 72
 Db 104 CCAGGCTTCAGATTCATCCAAACGACGACGAGGCTGGTGTGCATCTCTCGCGTAA 163
 QY 73 ValGluGlyLysArgPheAsnValGluLeuIleThrPheLeuAspLeuTyrArgTyrAsp 92
 Db 164 TCGCATCGCAGCTATCGCGTTTCCCATCATCGCGAAATCGACCTCTACAAATACGAC 223
 QY 93 ProTrpGluLeuProAlaMetAlaAlaIleGlyGluLysGluTyrTyrTyrValPro 112
 Db 224 CCCTGGGACCTACCGGATTCGCTTCCCTACGAGAGAAAGAGTGGTACTTCTTTCA 283
 QY 113 ArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrp 132

```
Db 284 CGGACCGGAATACCCGACCGTTCCAGGCGGACCGCGCGGGACCGGTTACTGG 343
Qy 133 LysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLys 152
Db 344 AAGCAACCGCGGGGATGAAGCCCATTTGCCAC-----CCCAACCGTTGGGATAAAA 397
Qy 153 LysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMet 172
Db 398 AAAGCTTTGGTGTTCACCGAGGNAAGCTTCGAAAGGGGACAGAGCAATGGATCATG 457
Qy 173 AsnGluTyrArgLeuProHisGluThrGluLysTyrGlnLysAlaGluSer--- 191
Db 458 CACGAGTATCTCTCGCGCATGTAGATCGCTCCGTTCCGCAAAAGAACACAGCTTAAGTTA 517
Qy 192 -----LeuCysArgValTyrLysArgProGly----- 200
Db 518 GATGATTGGGTGCTTTCGCGTATTACAAAGAGAGGCGACGATCGAAGAGTTCCAAACCA 577
Qy 201 -----ValGluAspHis----- 204
Db 578 AGCAGCGATGTGTGTAGCGGAAATGGAATCATCGGAGATCGAAGACAGGAGCCG 637
Qy 205 -----ProSerValPro-ArgSerLeuSerTh 213
Db 638 GAGATTCTGAAAGCGGAGGAGTGTCTTCTGCGCGCGGTTCGCGCGCCGCAAGGAAG 697
Qy 213 rArgHisGluAsnHisAsnSerSerThrSerSerArgLeuAlaLeuArgGlnGlnHis 233
Db 698 CGCGCGGTGAAGAGGATTACATGTACTTTCACCGCT---CGGATTCAATCCCGAAGCTG 754
Qy 233 eHisSerSerSer----- 238
Db 755 -CACACGACTCGAGCTGTTCCGAGCAGCTGGTATCGCGGAAATTCGCGAGCGAGGTGCA 813
Qy 239 -----AsnHisSerAs 242
Db 814 GAGCGAGCAAGTGAAGAGAGTGGAGAAAGCTCGAGTTTCGGTTTAAATACGTGA 873
Qy 242 pAsnAsnLeuAsnAsnAsnAsnAsnLeuAsnAsnLeuGluLysLeuSerThrGluTyrSe 262
Db 874 TGCCACTCTGAACACACAGCAACAGCTC----- 901
Qy 262 rGlyAspGlySerThrThrThrThrThrThrAsnSerAsnSerAspValThrIleAlaLe 282
Db 902 -----ACGACGCAATTCAGGCGCAATTAATCAGATGATGTCCCGCT 942
Qy 282 uAlaAsnGlnAsnIleTyrArgPro---MetProTyrAspThrSerAsnAsnThr----- 299
Db 943 GCAGGACATGTTTCATGTACTTGGCCCAACAGCCCTTCTGAACAAACCGATCCAGCCGCC 1002
Qy 300 -----LeuIleValSerThrArgAsnHis 307
Db 1003 CATTTCGATTTCACGGTAGCAGCAGACTCAAGGTCTATGATGTGCTGTGCGACCA 1062
Qy 307 sGlnAspAsp 310
Db 1063 CACAGACGAT 1072
```

RESULT 10

```
US-11-096-568A-26116
; Sequence 26116, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 26116
; LENGTH: 1189
```

```
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1189)
; OTHER INFORMATION: Ceres Seq. ID no. 13498621
US-11-096-568A-26116
```

```
Alignment Scores:
Pred. No.: 6,23e-39 Length: 1189
Score: 471.00 Matches: 103
Percent Similarity: 54.6% Conservative: 57
Best Local Similarity: 35.2% Mismatches: 85
Query Match: 23.2% Indels: 48
DB: 11 Gaps: 8
```

US-10-780-703-2 (1-379) x US-11-096-568A-26116 (1-1189)

```
Qy 43 AlaAspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThrGluGlu 62
Db 58 CGCGCGGAGCAGCAGCATCAGCTGGATCTCGCGGGTTCGGTTCCACCCGACGAGGAG 117
Qy 63 GluLeuIleGluPheTyrLeuArgLysValGluGlyLysArgPheAsnValGluLeu 82
Db 118 GAGCTGTGGAGTTCTACCTGAAGCAGGCGCGCGGAGGAGCTCAAGTTCGACATC 177
Qy 83 IleThrPheLeuAspLeuTyrArgTyrAspProTyrGluLeuProAlaMetAla---Ala 101
Db 178 ATCCCAAGGTGCACCTGTACCGGACGACCCCTGGGAGCTGCGAGCTGGCGCGGCTC 237
Qy 102 IleGlyGluLysGluTyrPheTyrValProArgAspArgLysTyrArg----- 118
Db 238 GTTGGCGAGCGGAGTGTACTTCTTCTGCGCGCGGAGCGCGCGCGCGCGCGCGCAAG 297
Qy 119 -----AsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrpLysAla 134
Db 298 CAGCAGCGCGGCTCCGCGCGCGCTTAGCGCACACGAGCGCGCGGTTCGGAAGGCC 357
Qy 135 ThrGlyAlaAspArgMetIleArg-----SerGluThrSerArgProIleGlyLeuLys 152
Db 358 ACGGGGTCGAGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCATCGGCTCAAG 417
Qy 153 LysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMet 172
Db 418 AAGACGCTCTCTACTACGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATG 477
Qy 173 AsnGluTyrArgLeuProHis-----HisGluThrGluLysTyrGlnLysAla 188
Db 478 AACGAGTACCGCTCCCGCGCGCTAGCAGCGCTAGCAGCGCGCGCGCGCGCGCGCGCGGTG 537
Qy 189 GluIleSerLeuCysArgValTyrLys-----Arg 198
Db 538 GAGCTGTGTGTGCAAGGTCTACCGCAAGCGCGTGTGCTCAAGGAGCTGGAGCAGCGG 597
Qy 199 ProGlyValGluAspHisProSerValProArgSerLeuSerThrArgHisAsnHis 218
Db 598 GTGGCCATGGAGGAG-----CTCGCGGCGCGCACCGCGCGCGCGCGCGCGCGCGCAC 648
Qy 219 -----AsnSerSerThrSerSerArgLeuAlaLeu----- 228
Db 649 AGCACCGGCTCCCGCGCGCTCATGCTCGTGGAGATCGTGTATGTCGCGCGTGGGC 708
Qy 229 -----ArgGlnGlnHisSerSerSer 237
Db 709 GCGCTCTCGTCTACCGTCAAGTGCATGAGCATGAGAGAGAGAGCGCGCGCGCGCGCGGANT 768
Qy 238 SerAsnHisSerAspAsnLeuAsnAsnAsnAsnAsnAsnAsnAsnLeuGluLysLeu 257
Db 769 CAGCAGCAGCAGGACCCCTCTCTGACGACGCTGGGAGCGCGCGCGCGCGCGCGCGCGTGG 828
Qy 258 SerThrGluTyrSerGlyAspGlySerThrThrThrThrThrThrThrThrThrThrThr 277
Db 829 TCTGCTTACTACGCGCGCTCTCAACTTCTTAACCAAAATCCACAACACAGCTTT 888
```


| | | | |
|-----|----|---|-----|
| 775 | Db | TTATCTGGAATCATACCAACCCACACAGGGGTTGGATCAAGTGCTGGATTCAACACT | 834 |
| 252 | Qy | AsnAsnIeuGluLysLeuSerThrGluTyrSerGlyAspGlySerThrThrThrThrThr | 271 |
| 835 | Db | GGGAATTGGATAGCAATGTCTCCAATCAATAACAACACCCCTCAACATAATTACATG | 894 |
| 272 | Qy | ThrAsnSerAsnSerAspValThrIleAlaLeuAlaAsn--GlnAsnIleTyrArgPro | 290 |
| 895 | Db | TTTCAGAAGAATCCCAATCGAACCCCTTCAGTCTCCAATCGAATGGAAAAATGATGAGCCA | 954 |

RESULT 13

```

US-11-096-568A-24275
; Sequence 24275, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 24275
; LENGTH: 1833
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1833)
; OTHER INFORMATION: Ceres Seq. ID no. 12424091
US-11-096-568A-24275

```

| Alignment Scores: | |
|------------------------|---------------|
| Pred. No.: | 2,32e-36 |
| Score: | 448.50 |
| Percent Similarity: | 46.2% |
| Best Local Similarity: | 32.7% |
| Query Match: | 22.1% |
| DB: | 11 |
| | Gaps: |
| | Indels: |
| | Mismatches: |
| | Conservative: |
| | Matches: |
| | Length: |

US-10-780-703-2 (1-379) x US-11-096-568A-24275 (1-1833)

| | | | |
|-----|--|---|-----|
| 51 | ValMetPro----- | GlyPheArgPheHisProThrGluGluGluLeuLeuGlu | 66 |
| | | | |
| 248 | GTGATGCCAACCTCCCGCGCGGGTTCGGCTTCACCCACGAGACGAGAGCTCATGGTG | 307 | |
| | | | |
| 67 | PheTyrLeuArgArgGlyValGluGlyValysArgPheAsnValGluLeuLeuThrPheLeu | 86 | |
| | | | |
| 308 | CACTACTCATGAGGAGCGCGCTCCATGCGGTCCGCTCCCATCATCATGCCAGGTC | 367 | |
| | | | |
| 87 | AspLeuTyrArgTyrAspProTyrPdpLeuProAlaMetAlaAlaIleGlyGluLeuysGlu | 106 | |
| | | | |
| 368 | AACATCTACCAAGTGCACACCCCTCGGGATCTCCCTGCCAAGGCATTGTTCCGCGACCAAGGAG | 427 | |
| | | | |
| 107 | TyrTyrPheTyrValProArgAspArgLysTyrArgAsnGlyAspArgProAsnArgVal | 126 | |
| | | | |
| 428 | TGGTTTTTCTTCAGCCCGCGGACCGCAAGTACCCCAACGGCGCCCGCCCCCAACCGCGCC | 487 | |
| | | | |
| 127 | ThrThrSerGlyTyrTyrPLeysAlaThrGlyValAspArgMetIleArgSerGluThrSer | 146 | |
| | | | |
| 488 | GCCGGGTCCGGGTACTGGAAGGCCACCGGCACCGGCAAGGCCATCTGTGCG--TCGTCC | 544 | |
| | | | |
| 147 | ArgPro----- | IleGlyLeuLeuysThrLeuValPhe | 157 |
| | | | |
| 545 | ACGCCACGACGACCGCGCGGCCAACATCGTCTGCTCGCGCTCAAGAAGCGCGCTGCTCTTC | 604 | |
| | | | |
| 158 | TyrSerGlyLysAlaProLysGlyThrArgThrSerTyrIleMetAsnGluTyrArgLeu | 177 | |
| | | | |
| 605 | TACGGCGGACGGCCGCCCAAGGACCAAGACGGACTGGATCATGCACGAGTACCGCTC | 664 | |
| | | | |
| 178 | Pro----- | HisHisGluThrGluLysTyrGlnLysAlaGluLeu | 190 |
| | | | |
| 665 | TCGGGCGCGGCGACGACGACTGCAAGGGGACGACCGAGCGCAGAGTGTCTTCCTCTCG | 724 | |
| | | | |

QY 53 ProGlyPheArg-PheHisProThrGluGluGluLeuLeuGluPheTyrLeuArgArgly 72
Db 112 CGGGGTTCCGGTTCCACCGGACGAGGAGTGGTGATGCATCTCTGCGCGG 171
QY 72 sValGluGlyLysArgPheAsnValGluLeuLeuLeuThrPheLeuAspLeuTyrArgTyrAs 92
Db 172 CTGCGCGCGCTGCCCATCGCGTCCCATCGCGGATCGACCTCTCAAGTTGCA 231
QY 92 pProTrpGluLeuProAlaMetAlaAlaLeuGluGluGluTyrPheTyrValPr 112
Db 232 CCCATGCGAGCTCCCAAGATGCGCTGTACGCGGAGAGGAGTGTACTCTCTCC 291
QY 112 oArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTr 132
Db 292 GCGGACCCGACGATACCCGACGGGTCCAGGCCCAACCGCGCGCGGTGGGTACTG 351
QY 132 pLysAlaThrGlyAlaAspArgMetLeuArgSerGluThrSerArgProLeuLeuLeu 152
Db 352 GAAGGCCACCGCGCTGACAAAGCCCTG-----GGCACGCCCAAGCGCTGCCCATCAA 405
QY 152 sLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpLeuMe 172
Db 406 GAAGCGCTCGTCTTCTACGCGGCAAGCGGCCCAAGGGCGAGAGCAACTGGATCAT 465
QY 172 tAnGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLysAlaGluLeuSer-- 191
Db 466 GCACGAGTACCGCTCGCGACGTCGCGCGCTCGCGCGCGCAAGAGAGAGCGCTCAGGTT 525
QY 192 -----LeuCyArgValTyrLysArgProGly 200
Db 526 CGATGACTGGTCTCTGCGCATCTACCAAGAGAGGCG 565

RESULT 15

US-11-096-568A-30488
; Sequence 30488, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096, 568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 30488
; LENGTH: 897
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(897)
; OTHER INFORMATION: Ceres Seq. ID no. 4961610
US-11-096-568A-30488

Alignment Scores:
Pred. No.: 2.57e-36 Length: 897
Score: 444.50 Matches: 92
Percent Similarity: 54.8% Conservative: 51
Best Local Similarity: 35.2% Mismatches: 93
Query Match: 21.9% Indels: 25
DB: 11 Gaps: 7

US-10-780-703-2 (1-379) x US-11-096-568A-30488 (1-897)

QY 53 ProGlyPheArgPheHisProThrGluGluGluLeuLeuGluPheTyrLeuArgArgly 72
Db 13 CCAGGATTCAGGTTTCATCCCACTGATGAGAACTGGTGGCTTACTATCTTGATAGAG 72
QY 73 ValGluGlyLysArgPheAsnValGluLeuLeuThrPheLeuAspLeuTyrArgTyrAsp 92
Db 73 GTCAACGGCCAGCCATTGAGCTCGAGATCATCCAGAGAGTTGATCTTTATAAATCGGAG 132
QY 93 ProTrpGluLeuProAlaMetAlaAlaLeu-----GlyGluLysGluTrpTyrPheTyr 110

Db 133 CCATGGGACTTGCCTGAAAAAGTCATTTTGGCGGAAACGACGATGATGGTACTTTTAC 192
QY 111 ValProArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGly 130
Db 193 AGCACAGGGGATAAGAGTATATCCAAATGGCTCTAGGACGAACCGTGCACGCGGGGT 252
QY 131 TyrTrpLysAlaThrGlyAlaAspArgMetLeuArgSerGluThrSerArgProLeuGly 150
Db 253 TACTGGAAGGCCACCGGAAAGATGCTACAGTAGAATCAAGAAGATGAAG---ATGGGA 309
QY 151 LeuLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrp 170
Db 310 ATGAAGAAGACACTGGTTTATTATAGAGGAAGGGCTCTCATGGCTTCGTACTAATTCG 369
QY 171 IleMetAsnGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLysAlaGluLeu 190
Db 370 GTCATGTCATGAATATCGTCTCACGACGCTCTCTCTCTCTCTCTG---AAGGAGTCGTAT 426
QY 191 SerLeuCyArgValTyrLysArgProGlyValGluAspHisProSerValProArgSer 210
Db 427 GCATTGTCCCGAGTGTTTAAGAAG-----AACATACAAATTCCAAGAGA 471
QY 211 LeuSerThrArgHisHisAsnHisAsnSerSerThrSerArgLeuAlaLeuArgGln 230
Db 472 AAAGGGGAAGAAGAAGACAGACAGACACTAGT-----GTAGGAAAAAGAGAG 525
QY 231 GlnGlnHisSerSerSerSerSerHisSerAspAsnAsnLeuAsnAsnAsnAsn 250
Db 526 GAAGAAAGAAAGGAG 585
QY 251 IleAsnAsnLeuGluLysLeuSerThrGluTyrSerGlyAspGlySerThrThrThr 270
Db 586 TTGAAAAGA-----GCATCCCGGAGAGACATCT 612
QY 271 ThrThrAsnSerAsnSerAspValThrIleAlaLeuAlaAsnGlnAsnIleTyrArgPro 290
Db 613 TCATCAGAGCTAACTCAAGGGTCTCTTTAGACGAAGCAACAGCTCATCCATTTTGTCT 672
QY 291 MetProTyrAspThrSer-----AsnAsnThrLeuIleValSerThrArgAsnHis 307
Db 673 CTTTCAATTTCTCATCTCTCTCTCTGACGATCATGATCATCTTTTCTCAAACTATTCTCAT 732
QY 308 Gln 308
Db 733 CAG 735

Search completed: April 6, 2006, 10:34:43
Job time : 588 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 6, 2006, 09:58:27 ; Search time 239 Seconds
(without alignments)
2818.811 Million cell updates/sec

Title: US-10-780-703-2

Perfect score: 2034

Sequence:

1 MAIVSSTTSIIPMSNQVNNN.....MNPVDPDGNDRDHTYNIPFK 379

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+p2n.model -DEV=xlh
-Q=/abses/ABSSWEB.spool/US10780703/runat_06042006_101342_6427/app_query.fasta_1
-DB=Issued Patents NA -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss805h -USER=US10780703 @CGN_1_193 @runat_06042006_101342_6427
-NCFU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/1/COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5/COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A/COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B/COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H/COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP.COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE.COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 509 | 25.0 | 1209 | 3 | US-09-614-408-6 |
| 2 | 498.5 | 24.5 | 1211 | 3 | US-09-614-408-1 |
| 3 | 498.5 | 24.5 | 1211 | 3 | US-09-614-981-1 |
| 4 | 494 | 24.3 | 1212 | 3 | US-09-533-029-3 |
| 5 | 480 | 23.6 | 1209 | 3 | US-09-533-029-37 |
| 6 | 462 | 22.7 | 1216 | 3 | US-09-533-029-121 |
| 7 | 455 | 22.4 | 914 | 3 | US-09-533-029-23 |
| 8 | 403 | 19.8 | 1287 | 3 | US-09-889-926-1 |
| 9 | 357 | 17.6 | 913 | 3 | US-09-533-029-73 |

| | | | | | | | |
|---|----|-------|------|---------|---|----------------------|--------------------|
| c | 10 | 330 | 16.2 | 1332 | 3 | US-09-533-029-115 | Sequence 115, Appl |
| | 11 | 293 | 14.4 | 2680 | 3 | US-09-614-408-7 | Sequence 7, Appli |
| | 12 | 239.5 | 11.8 | 1339 | 3 | US-09-533-029-99 | Sequence 99, Appl |
| | 13 | 131 | 6.4 | 1653 | 3 | US-09-248-796A-336 | Sequence 336, Appl |
| | 14 | 128 | 6.3 | 2094 | 3 | US-09-248-796A-1192 | Sequence 1192, A |
| | 15 | 125.5 | 6.2 | 934 | 3 | US-09-270-767-12547 | Sequence 12547, A |
| c | 16 | 122.5 | 6.2 | 1086 | 2 | US-08-415-751-47 | Sequence 47, Appl |
| | 17 | 122.5 | 6.0 | 1107 | 3 | US-09-614-221A-398 | Sequence 398, Appl |
| | 18 | 122.5 | 6.0 | 1974 | 3 | US-09-248-796A-11107 | Sequence 11107, A |
| | 19 | 122 | 6.0 | 2001 | 3 | US-09-487-558B-35 | Sequence 35, Appl |
| | 20 | 121.5 | 6.0 | 2190 | 3 | US-09-625-188-19 | Sequence 19, Appl |
| | 21 | 121 | 5.9 | 750 | 3 | US-09-248-796A-9460 | Sequence 9460, Ap |
| | 22 | 120 | 5.9 | 1431 | 3 | US-09-248-796A-5089 | Sequence 5089, Ap |
| | 23 | 118.5 | 5.8 | 2115 | 3 | US-09-614-221A-111 | Sequence 111, App |
| | 24 | 118.5 | 5.8 | 2115 | 3 | US-09-487-558B-217 | Sequence 217, App |
| | 25 | 116 | 5.7 | 285 | 3 | US-09-313-294A-2129 | Sequence 2129, Ap |
| | 26 | 115.5 | 5.7 | 1020 | 3 | US-09-248-796A-1997 | Sequence 1997, Ap |
| | 27 | 115.5 | 5.7 | 1230025 | 3 | US-09-198-452A-1 | Sequence 1, Appli |
| | 28 | 115.5 | 5.7 | 1230230 | 3 | US-09-438-185A-1 | Sequence 1, Appli |
| | 29 | 114.5 | 5.6 | 789 | 3 | US-09-248-796A-11671 | Sequence 11671, A |
| | 30 | 114.5 | 5.6 | 2574 | 3 | US-09-248-796A-6419 | Sequence 6419, Ap |
| | 31 | 113.5 | 5.6 | 723 | 3 | US-09-248-796A-13410 | Sequence 13410, A |
| | 32 | 113.5 | 5.6 | 867 | 3 | US-09-216-393B-340 | Sequence 340, App |
| | 33 | 113.5 | 5.6 | 867 | 3 | US-09-216-393B-342 | Sequence 342, App |
| | 34 | 113.5 | 5.6 | 1397 | 3 | US-09-216-393B-343 | Sequence 343, App |
| c | 35 | 113.5 | 5.6 | 1397 | 3 | US-09-216-393B-345 | Sequence 345, App |
| | 36 | 113.5 | 5.6 | 1614 | 3 | US-09-892-360-1 | Sequence 1, Appli |
| | 37 | 113.5 | 5.6 | 2710 | 3 | US-09-949-016-1497 | Sequence 1497, Ap |
| | 38 | 113.5 | 5.6 | 2730 | 3 | US-09-799-451-432 | Sequence 432, App |
| | 39 | 113.5 | 5.6 | 2730 | 3 | US-09-949-016-1130 | Sequence 1130, Ap |
| | 40 | 112.5 | 5.5 | 3456 | 3 | US-09-487-558B-133 | Sequence 133, App |
| | 41 | 111.5 | 5.5 | 65848 | 3 | US-09-949-016-13285 | Sequence 13285, A |
| c | 42 | 111 | 5.5 | 1125 | 3 | US-09-949-016-11171 | Sequence 11171, A |
| | 43 | 111 | 5.5 | 1800 | 3 | US-09-270-767-11171 | Sequence 11171, A |
| | 44 | 111 | 5.5 | 2277 | 3 | US-09-487-558B-223 | Sequence 223, App |
| | 45 | 110 | 5.4 | 2733 | 3 | US-08-997-685A-1 | Sequence 1, Appli |

ALIGNMENTS

RESULT 1

US-09-614-408-6
; Sequence 6, Application US/09614408
; Patent No. 6762347
; GENERAL INFORMATION:
; APPLICANT: GIOVANNONI, JAMES
; APPLICANT: TANKSLEY, STEVEN
; APPLICANT: VREBALOV, JULIA
; APPLICANT: NOENSIE, FREDERICK
; TITLE OF INVENTION: NOR GENE COMPOSITIONS AND METHODS FOR USE THEREOF
; FILE REFERENCE: TANK:213
; CURRENT APPLICATION NUMBER: US/09/614,408
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 60/143,357
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Tomato
US-09-614-408-6

Alignment Scores:
Pred. No.: 6.16e-48
Score: 509.00
Percent Similarity: 52.6%
Best Local Similarity: 35.0%
Query Match: 25.0%
DB: 3
Length: 1209
Matches: 113
Conservative: 57
Mismatches: 89
Indels: 64
Gaps: 10

US-10-780-703-2 (1-379) x US-09-614-408-6 (1-1209)


```

Db      678 ACAATATTCTGCGAAATAAGTTCGAATATATGACGATGGAATTATGAATAAACACGAACGAT 737
      ::::  |||||  |||  ::::  |||||  |||
Qy      248 -----AsnAsnAsnIleAsnAsnLeuGluLysLeuSerThrGlu----- 260
      |||||  |||||  |||||  |||||  |||||  |||||
Db      738 ATTATCAACAATAATAATAGATCCATTCCACAATAATCGTCAAGAGAACGATGCATGGA 797
      ::::  |||||  |||||  |||||  |||||  |||||
Qy      261 -----TySerGlyAspGlySerThrThrThrThrThrThrAsnSerAsnSerAsp 277
      ::::  |||||  |||||  |||||  |||||  |||||
Db      798 GGTGTTGATTTGGAATAACGACGAAGCAACAACAACAACAACAACCT----- 842
      ::::  |||||  |||||  |||||  |||||  |||||
Qy      278 ValThrIleAlaAlaAsnGlnAsnIleTyArgProMetProTyArgPThrSerAsn 297
      ::::  |||||  |||||  |||||  |||||  |||||
Db      843 -----ATTGATAGGAACCAATCTCTCCA 863
      ::::  |||||  |||||  |||||  |||||  |||||
Qy      298 AsnThrLeuIleValSer-ThrArgAsnHisGlnAspAspGluThrAlaIleValAs 317
      |||||  |||||  |||||  |||||  |||||  |||||
Db      864 AATACAAAAAGGTTTCCTTGTTGTGAGACAACACGAGGACGATGACTTAAC-----ATGAA 917
      |||||  |||||  |||||  |||||  |||||  |||||
Qy      317 pAspLeuGlnArgLeuValAsnTyGlnIleSerAspGlyAlaThrThrLeuMetProG 337
      ::::  |||||  |||||  |||||  |||||  |||||
Db      918 TAATATTTCGCAATTACAAATCATGAACAAAGTAGCTCCATTGCCAATTTCTCGAGCCA 977
      ::::  |||||  |||||  |||||  |||||  |||||
Qy      337 n 337
Db      978 G 978

RESULT 3
US-09-614-981-1
; Sequence 1, Application US/09614981
; Patent No. 6787687
; GENERAL INFORMATION:
; APPLICANT: GIOVANNONI, JAMES
; APPLICANT: TANKSLEY, STEVEN
; APPLICANT: PADMANABHAN, VEERARAGAVAN
; APPLICANT: RUEZINSKY, DIANE
; APPLICANT: VREBALOV, JULIA
; APPLICANT: WHITE, RUTH
; TITLE OF INVENTION: RIN GENE COMPOSITIONS AND METHODS FOR USE THEREOF
; FILE REFERENCE: TANK:214
; CURRENT APPLICATION NUMBER: US/09/614,981
; CURRENT FILING DATE: 2000-07-12
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1211
; TYPE: DNA
; ORGANISM: Tomato
US-09-614-981-1

Alignment Scores:
Pred. No.:          9.72e-47          Length:      1211
Score:             498.50             Matches:     116
Percent Similarity: 51.9%             Conservative: 61
Best Local Similarity: 34.0%          Mismatches:   98
Query Match:       24.5%              Indels:      66
DB:                3                  Gaps:       11

US-10-780-703-2 (1-379) x US-09-614-981-1 (1-1211)

Qy      37 ValGlnAsnGluAspGluAlaAspAspHisAspHisAsp---MetValMetProGlyPhe 55
      ::::  |||||  |||||  |||||  |||||  |||||
Db      36 ATGGAAGTACGATTCATCAACCGGGACAGCTCATCAGCTCAACTCCACCGCGGT 95
      ::::  |||||  |||||  |||||  |||||  |||||
Qy      56 ArgPheHisProThrGluGluLeuIleGluPheTyThrLeuArgArgLysValGluGly 75
      |||||  |||||  |||||  |||||  |||||  |||||
Db      96 CGATTCCACCCACGGAGCAAGAACTCATCTCCACTACCTCAAAAACACCAAGTCGCGCGC 155
      |||||  |||||  |||||  |||||  |||||  |||||
Qy      76 LysArgPheAsnValGluLeuIleThrPheLeuAspLeuTyArgTyArgPThrProGlu 95
      ::::  |||||  |||||  |||||  |||||  |||||
Db      156 GCTCCGATTCGGTGTGATATTATGCTGGAATATGATCTTTATATGTTTATCATCGGAA 215
      |||||  |||||  |||||  |||||  |||||  |||||
Qy      96 LeuProAlaMetAlaAlaIleGlyGluLysGluTrpTyThrPheTyValProArgAspArg 115

```

| | | | |
|----|-----|--|-----|
| Db | 216 | CTCCCTGCTAAGGCAATATTTCGGAGAGCAAGAAATGGTTCTTTTATTAGTCCCAAGAGATAGA | 275 |
| Qy | 116 | LVSTYrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyTrpIysAlaThr | 135 |
| Db | 276 | AAATATCTTAACGGGGAGGCCAAATCGGGCTGCAACATCGGGTTATTGGAAAGCTACC | 335 |
| Qy | 136 | GlyAlaAspArg---MetIleArgSerGluThrSerArgProIleGlyLeuIysIysThr | 154 |
| Db | 336 | GGAAACCCACAAAGCCGGTTTTTACTTTCGGTGGAAACAAAAAGTTGGGGTAAAAAAGGGC | 395 |
| Qy | 155 | LeuValPheTySerGlyIysAlaProIysGlyThrArgThrSerTrpIleMetAsnGlu | 174 |
| Db | 396 | CTCGTTTTTACGGCGGTAAACCAACCAAGGGTAAANAACTAATTGGATCATGATGAA | 455 |
| Qy | 175 | TyrArgLeuProHisHisGluThrGluIys----- | 184 |
| Db | 456 | TACAGAGTTGTAGAAAAATAAAACAATAACAAGCACCTTGGTTGTGATAATAATTGTTGCC | 515 |
| Qy | 185 | TyrGlnIysAlaGluIleSer-----LeuCysArgValTyIysArgPro | 199 |
| Db | 516 | AACAAAAAAGGATCTTTTCGAGCTAGATGATTTGGGTTTATCTCGAATTTTACAAGAAGAT | 575 |
| Qy | 200 | GlyValGluAspHisProSerValProArgSerLeuSerThrArgHisHis----- | 216 |
| Db | 576 | AACACACAA-----AGGTCATAGATGATTTCATGATATGTTGGGA | 617 |
| Qy | 217 | -----AsnHisAsnSerSerThrSerSerArgLeuAlaLeuArgGlnGlnHisHis | 234 |
| Db | 618 | TCGNATACCACAAATGTACCAAAATTCATATTACAAAGGAATAAAGCCTTCAAACTATGCT | 677 |
| Qy | 235 | Ser-----SerSerSerAsnHisSerAspAsnAsnLeuAsnAsn----- | 247 |
| Db | 678 | ACAATATTGCTCGAAAAATGAATCGAATATGTACGATGGAATTATGAAATAACACGAACGAT | 737 |
| Qy | 248 | -----AsnAsnAsnIleAsnAsnLeuGluIysLeuSerThrGlu----- | 260 |
| Db | 738 | ATTATCAACAAATAATAATAGATCCTCCACAATAATCGTCAAAGAGAACGATGCATGGA | 797 |
| Qy | 261 | -----TyrSerGlyAspGlySerThrThrThrThrThrThrAsnSerAsnSerAsp | 277 |
| Db | 798 | GGTTGTGATTGGAATACGACGAGCAACACACACCACT----- | 842 |
| Qy | 278 | ValThrIleAlaLeuAlaAsnGlnAsnIleTyrArgProMetProTyrAspThrSerAsn | 297 |
| Db | 843 | -----ATTGATAGGAACCACTTCTCCA | 863 |
| Qy | 298 | AsnThrLeuIleValSer-ThrArgAsnHisGlnAspAspGluThrAlaIleValAs | 317 |
| Db | 864 | ANTACAAAAAGGTTTCCTTGTGTGAGAACACAGAGACGATGGACTTAAC-----ATGAA | 917 |
| Qy | 317 | pAspLeuGlnArgValAsnTyrlGlnIleSerAspGlyAlaThrThrLeuMetProG | 337 |
| Db | 918 | TAATATTTCCGGAATTACAATCATGAACAAAGTAGTCCCATTCGCAATTTCTTGAGCCA | 977 |
| Qy | 337 | n | 337 |
| Db | 978 | G | 978 |

RESULT 4

US-09-533-029-3

; Sequence 3, Application US/09533029

; Patent No. 666446

; GENERAL INFORMATION:

; APPLICANT: Heard, Jacqueline

; APPLICANT: Broun, Pierre

; APPLICANT: Rischmann, Josee-Luis

; APPLICANT: Keddie, James

; APPLICANT: Pineda, Omaira

; APPLICANT: Adam, Luc

; APPLICANT: Samaha, Raymond

; APPLICANT: Zhang, James

; APPLICANT: Yu, Guo-Liang

RESULT 4
US-09-533-029-3
Sequence 3, Application US/09533029
; Patent No. 6654446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddie, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang

APPLICANT: Ratcliffe, Oliver
APPLICANT: Pilgrim, Marsha
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Reuber, Lynne
FILE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
FILE REFERENCE: MBI-010
CURRENT APPLICATION NUMBER: US/09/533,029
CURRENT FILING DATE: 2000-03-22
EARLIER APPLICATION NUMBER: 60/125,814
EARLIER FILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 121
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 1212
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: G759
US-09-533-029-3

Alignment Scores:
Pred. No.: 3,17e-46 Length: 1212
Score: 494.00 Matches: 100
Percent Similarity: 61.3% Conservative: 33
Best Local Similarity: 46.1% Mismatches: 64
Query Match: 24.3% Indels: 20
DB: 3 Gaps: 4

US-10-780-703-2 (1-379) x US-09-533-029-3 (1-1212)

| | | | |
|----|-----|--|-----|
| QY | 53 | ProGlyPheArgPheHisProThrGluGluLeuLeuPheTyrLeuArgArgLys | 72 |
| DB | 53 | CGGGTTCCGATTTTACCGCGGAGAGAGCTTATGTTCAATATCTCTGTAGAAA | 112 |
| QY | 73 | ValGluGlyLysArgPheAsnValGluLeuLeuThrPheLeuAspLeuTyrArgTyrAsp | 92 |
| DB | 113 | GCAGCTGGTTACGATTTCTCTCTCAGCTCATCGCGAAATAGATCTTTACAAATTCGAT | 172 |
| QY | 93 | ProThrGluLeuProAlaMetAlaAlaIleGlyGluLysGluTyrPheTyrValPro | 112 |
| DB | 173 | CCATGGTGGTTTACCAATAAAGCATATTATTGGAGAAAAAGATGGTATTTTATGTCCT | 232 |
| QY | 113 | ArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrp | 132 |
| DB | 233 | AGGATAGAAAATATCCAAACGGGTCAAGACCTAACCGGTTCCCGATCGGTTATTCG | 292 |
| QY | 133 | LysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLys | 152 |
| DB | 293 | AAAGCTACGGTACGGTAAATAATCTCGACGGAAGGACAAAGA--GTTGGTATTAA | 349 |
| QY | 153 | LysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMet | 172 |
| DB | 350 | AAAGCTTGGTGTGTACATCGGAAAGCTCTTAAAGTACTTAAACCAATTCGATCATG | 409 |
| QY | 173 | AsnGluTyrArgLeu-----ProHisHisGluThrGluLysTyrGlnLysAlaGluIle | 190 |
| DB | 410 | CATGAGTATCGTCTCATTTGAACTTCTCGTAGAAACGGAAGCACTAAGTTGATGATGG | 469 |
| QY | 191 | SerLeuCyArgValTyrLysArgProGlyValGluAspHisProSerValProArgSer | 210 |
| DB | 470 | GTTCTATGTCGAATATACAAAGAGCAATCAAGTGCACAAACAAACAGTTTACGATATGA | 529 |
| QY | 211 | LeuSerThrArgHisHisAsnHisAsnSerSerThrSerArgLeuAlaLeuArgGln | 230 |
| DB | 530 | ATCGCGAATGTAGAGAAATTCAGCAACAACGGTACTTCGTCCACG-----574 | |
| QY | 231 | GlnGlnHisHisSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer | 245 |
| DB | 575 | -----ACGTGCTCTTCTCTCACTTTTGAAGACGTTCTTGATTCGTTTCATCAA | 622 |
| QY | 246 | -----AsnAsnAsnAsnAsnIleAsnAsnLeuGluLysLeuSerThr | 259 |
| DB | 623 | GAGATCGAACACGAAATTTCCAGTTTCTTAAACCCAAACCGCATCTCGTCG | 673 |

RESULT 5
US-09-533-029-37
Sequence 37, Application US/09533029
Patent No. 666446
GENERAL INFORMATION:
APPLICANT: Heard, Jacqueline
APPLICANT: Broun, Pierre
APPLICANT: Riechmann, Jose-Luis
APPLICANT: Keddle, James
APPLICANT: Pineda, Omalra
APPLICANT: Adam, Luc
APPLICANT: Samaha, Raymond
APPLICANT: Zhang, James
APPLICANT: Yu, Guo-Liang
APPLICANT: Ratcliffe, Oliver
APPLICANT: Pilgrim, Marsha
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Reuber, Lynne
TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
FILE REFERENCE: MBI-010
CURRENT APPLICATION NUMBER: US/09/533,029
CURRENT FILING DATE: 2000-03-22
EARLIER APPLICATION NUMBER: 60/125,814
EARLIER FILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 121
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 37
LENGTH: 1409
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: G502
US-09-533-029-37

Alignment Scores:
Pred. No.: 1.6e-44 Length: 1409
Score: 480.00 Matches: 112
Percent Similarity: 51.5% Conservative: 39
Best Local Similarity: 38.2% Mismatches: 79
Query Match: 23.6% Indels: 64
DB: 3 Gaps: 7

US-10-780-703-2 (1-379) x US-09-533-029-37 (1-1409)

| | | | |
|----|-----|--|-----|
| QY | 53 | ProGlyPheArgPheHisProThrGluGluLeuLeuPheTyrLeuArgArgLys | 72 |
| DB | 248 | CCAGGTTTCCGATTTTACCCCTACCGATGAAGAGCTTGTCTATGCACCTATCTCTGCCGCAA | 307 |
| QY | 73 | ValGluGlyLysArgPheAsnValGluLeuLeuThrPheLeuAspLeuTyrArgTyrAsp | 92 |
| DB | 308 | TGTGCTCTCATGTCATCCGCTCCGATCATCGCTGAGATCGATCTCTCAAAATACGAT | 367 |
| QY | 93 | ProThrGluLeuProAlaMetAlaAlaIleGlyGluLysGluTyrPheTyrValPro | 112 |
| DB | 368 | CCATGGGAGCTTCTCGTTTAGCTTGTATGGTGAAGAGATGGTACTTCTCTCTCTCC | 427 |
| QY | 113 | ArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrp | 132 |
| DB | 428 | AGGACAGAAAAATATCCCAACGGTTCGGCTCTTAACCGCTCGCTGTTCTGTTACTGG | 487 |
| QY | 133 | LysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLys | 152 |
| DB | 488 | AAAGCTACCGGAGCTGATAAACCCGATC-----GGACTACTTAAACCGGTCGGAATTAAG | 541 |
| QY | 153 | LysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMet | 172 |
| DB | 542 | AAAGCTCTTGTCTTCTACCGCGCAAGCTCCAAAGGAGAGAGAAACCAATTTGGATCATG | 601 |
| QY | 173 | AsnGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLysAla-----188 | |
| DB | 602 | CACGAGTACCGTCTCGCCGACGTTTCACCGGTCGCTTCCGCAAGAGAAAGATAGTCTCAGG | 661 |

Qy 189 -----GluLeuSerLeuCyArGValTyrLysArgProGly----- 200
Db 662 CTGGATGATTGGTCTCTCGCGGATTATCAACAAAAAGGAGCTACCGAGAGCGCGGA 721
Qy 201 -----ValGluLeuSerLeuCyArGValTyrLysArgProGly----- 200
Db 722 CCACCGCTCGGTTGTTTACCGGACGAAATCATGGAGGAGAAAGCGGAGTGCACGAG 781
Qy 209 qSerLeuSerThrArgHisAsnHisAsn---SerSerThrSerSerArgLeuAlaLe 228
Db 782 ATGGTTATGCTCGCGCCGCAACAGCAAGTGGTTCGGTATTTTCGACACGTCGAT 841
Qy 228 uArgGlnGlnHisSerSerSerSerSerSerSerSerSerSerSerSerSerSer 246
Db 842 TCGGTGCGGAAGCTG-CATACTACGGATTTCGAGTTGCTCGGAGCAGGTGGTTCGCGCGGA 900
Qy 247 -----AsnAsnAsn 249
Db 901 GTTCACGAGCGAGGTTTCAGAGCGAGCCCAAGTGGAAGATTGGTTCGCGCGGTAAGTAATGA 960
Qy 249 nAsnLeuAsnLeuGlu-----LysLeuSerThrGluTy 261
Db 961 CAATAACATACCTCTGTTTGGGTTTAAATCATTCATGTCGCCACCGTGGATTAACGCGTT 1020
Qy 261 rSerGlyAspGlySerThrThrThrThrThrThrThrThrThrThrThrThrThr 281
Db 1021 TCGAGGAGGAGGAGT-----AGTAATCAGATGTTTCC 1053
Qy 281 aLeuAlaAsnGlnAsnLeuTyrArgProMetProTyr 293
Db 1054 GCTACAGGATATGTTTCATGTATCATGCAGAGACGCTTAC 1090

RESULT 6

US-09-533-029-121
; Sequence 121, Application US/09533029
; Patent No. 6664446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddie, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 121
; LENGTH: 1216
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G501
US-09-533-029-121

| Alignment Scores: | | |
|------------------------|----------|------------------|
| Prod. No.: | 1,42e-42 | Length: 1216 |
| Score: | 462.00 | Matches: 96 |
| Percent Similarity: | 59.8% | Conservative: 38 |
| Best Local Similarity: | 42.9% | Mismatches: 66 |
| Query Match: | 22.7% | Indels: 24 |
| DB: | 3 | Gaps: 6 |

US-10-780-703-2 (1-379) x US-09-533-029-121 (1-1216)
Qy 39 AsnGluAspGlu-----AlaAspAspHisAspMet----- 50
Db 17 AACGAAGAAGATTTTTCAGAGCAACAGCAAGAAGAAGATGAAGTCGAGCTA 76
Qy 51 ---ValMetProGlyPheArgPheHisProThrGluGluLeuLeuGluPheTyrLeu 69
Db 77 AATTTACAGCTGGTTCGATTCATCCACGAGCGAGGAGCTTGTGAATTTCTACTTG 136
Qy 70 ArgArgLysValGluGlyLysArgPheAsnValGluLeuLeuThrPheLeuAspLeuTyr 89
Db 137 TGCAGAAATGTCTCTCCGAGCAGATCTCGGCTCCGCTTATCCGAGATGTATCTCTAC 196
Qy 90 ArgTyrAspProThrGluLeuProAlaMetAlaAlaLeuGlyLysGluTyrPhe 109
Db 197 AGTTCAATCTTGGGAGCTTCAGAGATGTCTCTACGAGAGAAAGATGTGTACTTC 256
Qy 110 TyrValProArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrSer 129
Db 257 TTCTACCTAGAGATCGAAATACCCAAACGGTTCGGTCTTAACCGGCGAGGAAACC 316
Qy 130 GlyTyrTrpLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIle 149
Db 317 GGTATTGAAAGCTACCGAGCAGATAAACCGATTGGTAAA-----CCGAAGACGTTG 370
Qy 150 GlyLeuLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSer 169
Db 371 GGTATCAAGAAAGCACTCGTCTTCTACGAGGAAAGCTCCAAAGGGGATTAAAGCAAT 430
Qy 170 TrpIleMetAsnGluTyrArgLeuProHis-----HisGluThr 182
Db 431 TCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 490
Qy 183 GluLysTyrGlnLysAlaGluLeuSerLeuCyArgValTyrLysArgProGlyValGlu 202
Db 491 AACAACTACGACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 550
Qy 203 Asp-----HisProSerValProArgSerLeuSerThrArgHisAsnHisAsnSer 220
Db 551 GAGAAGTATTTTCCCGCGGATGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 610
Qy 221 Ser-----ThrSerSerArgLeuAlaLeuArgGlnGlnGlnHisHisSerSer 236
Db 611 TCACCTTTTGATACATCAGACTTACCTACCCACATTCGAGAGGATGATTCAGGAGC 670
Qy 237 SerSerAsnHis 240
Db 671 TCAGGTGGTCAAC 682

RESULT 7

US-09-533-029-23
; Sequence 23, Application US/09533029
; Patent No. 6664446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddie, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22

EARLIER APPLICATION NUMBER: 60/125,814

EARLIER FILING DATE: 1999-03-23

NUMBER OF SEQ ID NOS: 121

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 23

LENGTH: 914

TYPE: DNA

ORGANISM: Arabidopsis thaliana

FEATURE:

OTHER INFORMATION: G503

US-09-533-029-23

Alignment Scores:

Pred. No.: 5,648-42 Length: 914

Score: 455.00 Matches: 105

Percent Similarity: 50.0% Conservative: 30

Best Local Similarity: 38.9% Mismatches: 70

Query Match: 22.4% Indels: 65

DB: 3 Gaps: 8

US-10-780-703-2 (1-379) x US-09-533-029-23 (1-914)

```
Qy 53 ProGlyPheArgPheHisProThrGluGluGluLeuLeuLeuGluPheTyrLeuArgLys 72
Db 110 CCAGGGTTCGGATTTCATCTACCGACGAGAACTCATGTTTACTATCTCCGAAACCA 169
Qy 73 ValGluGlyLysArgPheAsnValGluLeuLeuLeuThrPheLeuAspLeuTyrArgTyr 92
Db 170 ACCATGCTCTAAACCATCGCCCTCTCCATCATCCAGAGTTGATATCTACAAATTCGAC 229
Qy 93 ProTTPGluLeuProAlaMetAlaLeuLeuGlyGluLeuTyrPheTyrValPro 112
Db 230 CATGGCAATACCCGAGAAAACAGAGTTGGAGAAAATGAGTGGTATTTCTCAGCCCT 289
Qy 113 ArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrp 132
Db 290 AGAAGAAGAAATATCCAAACCGAGTCAGACCAACCGGCGAGCTGTTCCGGTTATTGG 349
Qy 133 LysAlaThrGlyAlaAspArgMetLeuArgSerGluThrSerArgProLleGlyLeuLys 152
Db 350 AAAGCAACCGGTACAGACAAAGCCATTACAGCGGTTCGAGTAAC---GTAGGTGTCAAG 406
Qy 153 LysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpMet 172
Db 407 AAAGTCTCGTCTTCTACAAAGGTAGACCTCTAAAGGAATCAAAACTGACTGGATCATG 466
Qy 173 AsnGluTyrArgLeuProHisHisGluThrGluLys----- 184
Db 467 CATGATATCGTCTC-----CATGATTCAGTAAAGCATCAACGAACGTAGCGGATCT 520
Qy 185 TyrGlnLysAlaGluLeuSerLeuCysArgValTyrLysArgProGly----- 200
Db 521 ATGAGGTTAGATGAATGGGTACTATGTAGGATATACAAAGAGAGGACCAAGTAAGCTT 580
Qy 201 -----ValGluAspHisProSerVal--- 207
Db 581 CTGAATGAGCAAGAGGGTTTTCATGGACGAAGTACTAATGGAGGATGAGACCAAGTGT 640
Qy 207 ----- 207
Db 641 ATTAAGCAAGCAGAGAGAAATGATGAAGATATATGATGATGATGATGATGATGATGAT 700
Qy 208 ProArgSerLeuSerThrArgHis-----His 216
Db 701 CCAAGGACGTGTTCCGTGCTCATTTGTTGGAATGATGATGATGATGATGATGATGATGAT 760
Qy 217 AsnHisAsnSerSerThrSerArgLeuAlaLeuArgGlnGlnHisHisSerSer 236
Db 761 ATTGATAATTTTAGTCAGTTCGATCATCTT-----CATCAACCTGATTCGGAG 808
Qy 237 SerSerAsnHisSerAspAsnAsnLeuAsnAsnAsnAsnAsnLeuGluLys 256
Db 809 TCTAGTTGTTCCGGGATCTACAGTTTAAACCAAGACGAGATCTTAAACCATCATCTGTCNA 868
```

Qy 256 sLeuSerThrGluTyrSerGlyAspGly 265
Db 869 GCTATGTTTAAAGTTTATGTTT---GATGGG 893

RESULT 8

US-09-889-926-1

Sequence 1, Application US/09889926

Patent No. 6844486

GENERAL INFORMATION:

APPLICANT: Xie, Qi

APPLICANT: Chua, Nam-Hai

APPLICANT: Institute of Molecular Agrobiolgy, The National U

TITLE OF INVENTION: NAC1 - A PLANT GENE ENCODING A TRANSCRIPTION FACTOR

TITLE OF INVENTION: INVOLVED IN COTYLEDON AND LATERAL ROOT DEVELOPMENT

FILE REFERENCE: 2248-115

CURRENT APPLICATION NUMBER: US/09/889,926

CURRENT FILING DATE: 2001-08-18

NUMBER OF SEQ ID NOS: 2

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1

LENGTH: 1287

TYPE: DNA

ORGANISM: Arabidopsis thaliana

FEATURE:

NAME/KEY: CDS

LOCATION: (89)...(1060)

US-09-889-926-1

Alignment Scores:

Pred. No.: 8,356-36 Length: 1287

Score: 403.00 Matches: 116

Percent Similarity: 46.8% Conservative: 67

Best Local Similarity: 29.5% Mismatches: 144

Query Match: 19.8% Indels: 67

DB: 3 Gaps: 16

US-10-780-703-2 (1-379) x US-09-889-926-1 (1-1287)

```
Qy 12 ProMetSerAsnGlnValAsnAsnGluLysGlyIleGluAspAsnHisArgGly 31
Db 41 CCATCCCACTAATCAACACTAAACCTAGAAAA----- 73
Qy 32 GlyGlnGluSerHisValGlnAsnGluAspGluAlaAspAspHisAspMetVal 51
Db 74 AAAAGGATCAATCATGGAGACGAGAGAGATGAAGGAAAGTAGTATAGCATGGTG 133
Qy 52 -----MetProGlyPheArgPheHisProThrGluGluGluLeuLeu---Glu 66
Db 134 GAGGCAAGTTGCTCCCGGATTCAGATTTCCACCGAAGGACGATGCTTCTGCGAT 193
Qy 67 PheTyrLeuArgLysValGluGlyLysArgPheAsnValGluLeuLeuThrPheLeu 86
Db 194 TACTTGATGAGACGATCGCTTCACAATAATCATCCACCCTCTTGTCTGATCCCAAGTC 253
Qy 87 AspLeuTyrArgTyrAspProTTPGluLeuProAlaMetAlaAlaLeuGlyLysGlu 106
Db 254 GATCTCAACAAAGTGTGAGCCTTGGGACATCCCAAAATGCGATCGCTGGGCGGAGGAT 313
Qy 107 TrpTyrPheTyrValProArgAspArgLysTyrArgAsnGlyAspArgProAsnArgVal 126
Db 314 TGGTATTTCTACGCCAAGAGACCGAAAAATACCGCGCGGGCTGAGAACTAACCGAGCA 373
Qy 127 ThrThrSerGlyTyrTrpLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSer 146
Db 374 ACGGCCACCGGATATTGGAAGACCCACCGCAAGACAGAACCATTTCTAAGAAAG---GGT 430
Qy 147 ArgProLleGlyLeuLysLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThr 166
Db 431 AAGCTAGTTGGGATGAGGAAGACATTTGTTTTCTATCAAGGTCGAGCTCTCGAGCCCT 490
Qy 167 ArgThrSerTrpIleMetAsnGluTyrArgLeu-----ProHisHis 180
```

```
Db 491 AAAACCGATTGGGTATCCGATCCAGCAATTCGTTCCAGGATCTCATCATCTCCCAATCAT 550
QY 181 GluThrGluLysTyrglnLysAlaGluLeuSerLeuCyshArgValTyrlsArg----- 198
Db 551 TCCTGAGCTCT---CCAAAGGAAGACTGGGTCTTGTAGGGTATTCCATAAGATAACG 607
QY 199 ProGlyVal-----GluAspHis-----ProSer 206
Db 608 GAAGGAGTTATATCTAGACACAACATGGGAAGCTGTTTTGATGAGACAGCTCTCGCATCG 667
QY 207 ValProArgSerLeuSerThrArgHisAsnHisAsnSerSerThrSerSerArgLeu 226
Db 668 CTCCTCT---CCACTGAGTGATCTTATCATCACTTTGACCAAGAACCTCTCTTATCTC 724
QY 227 AlaLeuArgGlnGlnHisSerSerSerSerSerSerSerSerSerSerSerSerSer 244
Db 725 AGT-----GATGATCATCTATCATCATCATCATCATCATCATCATCATCATCATCAT 775
QY 245 LeuAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsn 264
Db 776 TTGTCACAGAACCACAACTTTAAAC----- 799
QY 265 GlySerThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 284
Db 800 -----TCGAACCTTAACCACTCACTCTCTGAACCTCAAGATTCCATGCAAGAAC 847
QY 285 GlnAsnIleTyrgProMetProTyrgProMetProTyrgProMetProTyrgProMet 304
Db 848 CTTAAC-----CCCTGTTTACTGGTGGTTCAGCCCTCAGCCAGCTCACAGGCTC 898
QY 305 ArgAsnHisGlnAspAspGluThrAlaIleValAspAspLeuGlnArgLeuValAsn 324
Db 899 GACTCATCTGTTCTTCAGATCAGATGTTCTCAGAGCTCTACTCAGTCACTCACTAAG 958
QY 325 TyGlnIleSerAspGlyAlaThrLeuMetProGlnThrGlnAlaAlaLeuAlaMet 344
Db 959 ATTGATGAAGCCCTCGGGCTTAAGAATCATCAGAGTTATGGAAGAGGTAGCTCGAGAGC 1018
QY 345 AsnMetIleProAlaGlyThrIleProAsnAsnAlaLeuTrpAspMetTrpAsnProIle 364
Db 1019 CTCCTACCGACATCGGT---ATTCCAGCACT-----GTTTGAATG-CTG 1062
QY 365 ValProAspGlyAsnArgAspHisTyrgThrAsnIlePro 377
Db 1063 ATGATCGAGTGTAAACGAGATTACTATTGCTATATTCCT 1101

RESULT 9
US-09-533-029-73
; Sequence 73, Application US/09533029
; Patent No. 666446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddie, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; EARLIER FILING DATE: 2000-03-22
; EARLIER FILING DATE: 60/125,814
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 73
```

```
; LENGTH: 913
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G511
US-09-533-029-73

Alignment Scores:
Pred. No.: 8 45e-31 Length: 913
Score: 357.00 Matches: 75
Percent Similarity: 61.4% Conservative: 27
Best Local Similarity: 45.2% Mismatches: 46
Query Match: 17.6% Indels: 18
DB: 3 Gaps: 5

US-10-780-703-2 (1-379) x US-09-533-029-73 (1-913)
QY 49 AspMetValMetProGlyPheArgPheHisProThrGluGluGluLeuIleGluPheTyr 68
Db 37 GATGAGTCAACAATCGGGTTTCGCTTCTATCCACGAAAGAACTGGTTTCGTTCTCTAC 96
QY 69 LeuArgArgLysValGluGlyLysArgPheAsnVal-----GluLeuIleThrPheLeu 86
Db 97 CTACGAAACAGCTCGAAGAAAGAGTGAATCAATGATCGTGTCTATCCCGTACTT 156
QY 87 AspLeuTyrgTyrgAspProTrpGluLeuProAlaMetAlaIle-----Gly 103
Db 157 GAGGCTTTGAGTTCGAGCTAGTCACTCTCCAAATGTTGCTGGAGTGAGATGTCGAGGA 216
QY 104 GluLysGlu---TrpTyrgPheTyrgValProArgAspArgLysTyrgArgAsnGlyAspArg 122
Db 217 GACGCTGAGCAATGGTTCCTCTTCGTGTCACGACAAAGAACGCCAAGCAAGAGGAGGAGA 276
QY 123 ProAsnArgValThrThrSerGlyTyrgTrpLysAlaThrGlyAlaAspArgMetIleArg 142
Db 277 CCGAGTAGAACTACTCGTTTCAGGATACTGGAAGCAACTGGATCCTCGTCCAGTCTTT 336
QY 143 SerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTyrgSerGlyLysAla 162
Db 337 TCCAAAGACACAAATAATGATTGGAGCAAGAAACTATGTTTCTACACTGGAAGACA 396
QY 163 ProLysGlyThrArgThrSerTrpIleMetAsnGluTyrgArgLeuProHisHisGluThr 182
Db 397 CCCACAGGAAGAAATAAATGGAATAATGAATGAGTAC-----CACGCCGTT 444
QY 183 GluLysTyrgLysAla-----GluLeuSerLeuCyshArg 194
Db 445 GACGAAACAGTCAACGCTTCCCAATCCCTAAGCTGAGAGCTGAGTTTATGTCGA 504
QY 195 ValTyrgLysArgProGly 200
Db 505 GTCTACATAACACAGGA 522

RESULT 10
US-09-533-029-115/c
; Sequence 115, Application US/09533029
; Patent No. 666446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddie, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
```

```
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 115
; LENGTH: 1332
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G521
US-09-533-029-115

Alignment Scores:
Pred. No.: 1,866-27 Length: 1332
Score: 330.00 Matches: 60
Percent Similarity: 55.7% Conservative: 28
Best Local Similarity: 38.0% Mismatches: 36
Query Match: 16.2% Indels: 34
DB: 3 Gaps: 2

US-10-780-703-2 (1-379) x US-09-533-029-115 (1-1332)
Qy 53 ProGlyPheArgPheHisProThrGluGluLeuLeuGluPheTyrLeuArgArgLys 72
Db 1311 CCTGTTTTCGGTTTCATCCCGACTGATGTTGAACCTTGAGATACCTACTTGAGAGGANA 1252
Qy 73 ValGluGlyLysArgPheAsnValGluLeuLeuThrPheLeuAspLeuTyrArgTyrAsp 92
Db 1251 GTATGCGGAAAAGTCCCAAGTTGATGCCATTGCTGAGTGCACATTACCAAGTTCGAA 1192
Qy 93 ProTrpGluLeuPro----- 97
Db 1191 CCTCTGATTATACCCGGTAACACTCACTCTTTTCAGAAAATCTTTGTTCTGTTTCATGCT 1132
Qy 98 -----AlaMet 99
Db 1131 TTCATTATTATTCCTGATTACTTAACAGAGGTTTCTTACATTGTACAGACAAATCATGT 1072
Qy 100 AlaAlaIleGlyLysGluTrpTyrPheTyrValProArgAspArgLysTyrArgAsn 119
Db 1071 CTAGGACTGGAGATCTTAAGTGGTACTTCTCTGTCGAAGAGAAAAGAGATATCCAAA 1012
Qy 120 GlyAspArgProAsnArgValThrThrSerGlyTyrTrpLysAlaThrGlyAlaAspArg 139
Db 1011 GCGCGTAAGGCAACCGCTCTACTGAATGTGTTACTTGAAGACCAACAGGGGAGACAGA 952
Qy 140 MetIleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTyrSer 159
Db 951 GATGTT---TCTATAATGACGAAGTCACTGCGGAAGATAGAACTCTGATTTATCACTAC 895
Qy 160 GlyLysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeu 177
Db 894 GGGAAAATACCTCGCGGTGATCGGACTGATGGTCAATACATAGATATAGACTT 841

RESULT 11
US-09-614-408-7
; Sequence 7, Application US/09614408
; Patent No. 6762347
; GENERAL INFORMATION:
; APPLICANT: GIOVANNONI, JAMES
; APPLICANT: TANKSLEY, STEVEN
; APPLICANT: VREBALOV, JULIA
; APPLICANT: NOENSIE, FREDERICK
; TITLE OF INVENTION: NOR GENE COMPOSITIONS AND METHODS FOR USE THEREOF
; FILE REFERENCE: TANK:213
; CURRENT APPLICATION NUMBER: US/09/614,408
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 60/143,357
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 7
; LENGTH: 2680
; TYPE: DNA
; ORGANISM: Tomato
US-09-614-408-7

Alignment Scores:
Pred. No.: 9,55e-23 Length: 2680
Score: 293.00 Matches: 75
Percent Similarity: 29.4% Conservative: 25
Best Local Similarity: 22.1% Mismatches: 30
Query Match: 14.4% Indels: 211
DB: 3 Gaps: 2

US-10-780-703-2 (1-379) x US-09-614-408-7 (1-2680)
Qy 53 ProGlyPheArgPheHisProThrGluGluLeuLeuGluPheTyrLeuArgArgLys 72
Db 87 CCGGGGTTTCGATTTCCACCCGACGCGAAGAACTCATCGTCCACTACCTCAAAAACGA 146
Qy 73 ValGluGlyLysArgPheAsnValGluLeuLeuThrPheLeuAspLeuTyrArgTyrAsp 92
Db 147 GTCGCCGCGCTCCGATTCGGTGGATATTATTCGGTGAATTCATCTTTAAGTTGAT 206
Qy 93 ProTrpGluLeuPro----- 97
Db 207 CC-TGGGACTCCCTGGTACTATTTTCCACCACACTATATATTTCTTGCCCTATAACTT 265
Qy 97 ----- 97
Db 266 ATATATAGGGGAAAAGATCGGAGTCGAGTCAGCGATGAACAATATTGTGCTAAATTTT 325
Qy 97 ----- 97
Db 326 TAAATATGCAATAGATTGGTGACGAATTCGTTGCTAAATTTTCTAGTGATAAATTA 385
Qy 97 ----- 97
Db 386 ATATTTTCCCTTTTAACTCTCATGTTTATATCATAAAGTTTCTATGACCAACTTA 445
Qy 97 ----- 97
Db 446 TAAAGATTGGAACTCGATCAATTTTTTTTTTAGAATGAATGAACCTTATGTTATATAGT 505
Qy 97 ----- 97
Db 506 GATATTTAAATGCTTTTTTATATTTTCAAAGATATCCACGATACCGTGTAAAAGTGA 565
Qy 97 ----- 97
Db 566 ATTTGCAAAAAAATGCTAGTACCTTTTAAATTTTATTGTTAGATAAATTTAGATTTT 625
Qy 97 ----- 97
Db 626 AATTTGAATTTGTTTAAATTTAAATTTCTGAATCGTATAATATTTTAAATTTCTATT 685
Qy 97 ----- 97
Db 686 TTGAGTTTTTTTTTGGAGGGTGCTTAAAAAGTAGTATTTCACAAATATAAAGTAGTGGA 745
Qy 97 ----- 97
Db 746 AACATAAGTAGTGAGCCCAATAATTTATTTTAAAAAATATATATAAATCTATTGTTA 805
Qy 98 -----AlaMetAlaIleGly 103
Db 806 AGTTTAAATTCGAATTATCTTCTTATCATGTGTTTAAACGCGCTTAAGGCAATATTCGA 865
Qy 104 GluLysGluTyrTrpTyrPheTyrValProArgAspArgLysTyrArgAsnGlyAspArgPro 123
Db 866 GAGCAAGAATGTTCTTTTTCCTCCCAAGAGATAGAAAATATCTCTTAACGGGCGGAGGCA 925
Qy 124 AsnArgValThrThrSerGlyTyrTriPysAlaThrGlyAlaAspArg---MetIleArg 142
```

Db 926 AATCGGGCTGCAACATCGGTTATTGGAAGCTACCGAACCGACCGCGTTTACT 985
Qy 143 SerGluThrSerArgProIleGlyLeuLysThrLeuValPheTyrSerGlyValAla 162
Db 986 TCCGGTGAACACAAAGGTTGGGGTAAAAAAGCGCTCGGTTTTTACGGCGGTAAACCA 1045
Qy 163 ProlysGlyThrArgThrSerTriPleMetAsnGluTyrArgLeuProHisHisGluThr 182
Db 1046 CCAAAAGGGGTAAAACTAATGGATGTCGATGATACAGAGTTGTAGAAAATAAACA 1105

RESULT 12

US-09-533-029-99

; Sequence 99, Application US/09533029

; Patent No. 6664446

; GENERAL INFORMATION:

; APPLICANT: Heard, Jacqueline

; APPLICANT: Broun, Pierre

; APPLICANT: Riechmann, Jose-Luis

; APPLICANT: Keddle, James

; APPLICANT: Pineda, Omaira

; APPLICANT: Adam, Luc

; APPLICANT: Samaha, Raymond

; APPLICANT: Zhang, James

; APPLICANT: Yu, Guo-Liang

; APPLICANT: Ratcliffe, Oliver

; APPLICANT: Pilgrim, Mareha

; APPLICANT: Jiang, Cai-Zhong

; APPLICANT: Reuber, Lynne

; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES

; FILE REFERENCE: MBI-010

; CURRENT APPLICATION NUMBER: US/09/533,029

; CURRENT FILING DATE: 2000-03-22

; EARLIER APPLICATION NUMBER: 60/125,814

; EARLIER FILING DATE: 1999-03-23

; NUMBER OF SEQ ID NOS: 121

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 99

; LENGTH: 1339

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; FEATURE:

; OTHER INFORMATION: G515

US-09-533-029-99

Alignment Scores:

Pred. No.: 3,94e-17 Length: 1339

Score: 239.50 Matches: 82

Percent Similarity: 45.7% Conservative: 50

Best Local Similarity: 28.4% Mismatches: 100

Query Match: 11.8% Indels: 58

DB: 3 Gaps: 11

US-10-780-703-2 (1-379) x US-09-533-029-99 (1-1339)

Qy 45 AspHisAspHisAspMetValMetPro---GlyPheArgPheHisProThrGluGlu 63
Db 139 GATCGCTGCAGAGAAATGGAAACTCCTGTGGTTTAAAGATTCTCTCCGACCGAGGAG 198
Qy 64 LeuIleGluPheTyrLeu-ArgArgLys-----ValGluGlyLysArgPheAsnValG1 81
Db 199 ATCGTCTCGATTACCTTTGGCCGAAAAATTCGATAGACACGACGCCATGTCGATCGA 258
Qy 81 uLeuIleThrPheLeuAspLeuTyrArgTyrAspProTyrGluLeuProAlaMetAlaA1 101
Db 259 TTCATTAAACACA-GTCCCTGTCTGTAGACTCGATCTTGGGAGTTACCTTGCCAGTCAAG 317
Qy 101 alleGlyLysGlu-----TrrTyrPheTyrValProArgAspArgLysTyrArgAs 119
Db 318 GATCAAACTGAAGATGTGGTCTGTTCTTCAGACCTTAAGGAGAACAAATATGGCAG 377
Qy 119 nGlyAspArgProAsnArgValThrThrSerGlyTyr-TripLysAlaThrGlyAlaAspAr 139

Db 378 AGGTGATCAGCAGATGAGAAAAACGAAATCTGGGTTTGGGAAGAGTACTGCGACACAAA 437
Qy 139 gMetIleArgSerGluThrSerArgProIleGlyLeuLysThrLeuValPheTyrSe 159
Db 438 GCCTATC---ATCGTAATCGCCAAACAGATCGGTGAGAAAAGATTTGATGTTTACAC 494
Qy 159 rGlyLysAlaProLysGlyThrArgThrSerTriPleMetAsnGluTyrArgLeuProHi 179
Db 495 GAGTAAGGAATCCAAA-----TCCGATTGGTTATACACGAGTAC-----CA 536
Qy 179 sHisGluThrGluLysTyrGlnLysAlaGluLysLeuSerLeuCysArgValTyrLysArgPr 199
Db 537 CGGTTCTCTCATAAACAGATGATGATACATACACTCTGTAAAGTTATGTTTAATGG 596
Qy 199 oGlyValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsnHisAs 219
Db 597 TGGCATG-----AGAGAGAAGTCTTCTCTC 620
Qy 219 nSerSerThrSerArgLeuAlaLeuArgGlnGlnHisHisSerSerSerSerAs 239
Db 621 TTCTCTCTCTCTCTCTGGTGTAGTGAATGAGCAGAGTCGTGTCGACTCTTTAATC-- 678
Qy 239 nHisSerAspAsnAsnLeuAsnAsnAsnLeuAsnAsnLeuLysLeuSerTh 259
Db 679 -----CCTCAGCTTGTCACAAATCTCGAGGATCCTCACTTCACAGAGAAGATCC 728
Qy 259 rGluTyrSerGlyAspGlySerThrThrThrThrThrThrThrThrThrThrThr 279
Db 729 AGTCAGTTTGGTGAT----- 744
Qy 279 rIleAlaLeuAlaAsnGlnAsnIleTyrArgProMetProTyrAspThrSerAsnAsnTh 299
Db 745 -----CTGTCGCAAGAAGTCCCAATCGAGGATCGTAACACTGAC 782
Qy 299 r-----LeuIleValSerThrArgAsnHisGlnAspAspAspGluThrAlaIleValAs 317
Db 783 CGAGGAATGGTAAATGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 833
Qy 317 pAsp-LeuGlnArgLeuValAsn 324
Db 834 GGATGCTATACCGATTGAGGAAT 856

RESULT 13

US-09-248-796A-336

; Sequence 336, Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBI

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 336

; LENGTH: 1653

; TYPE: DNA

; ORGANISM: Candida albicans

US-09-248-796A-336

Alignment Scores:

Pred. No.: 0.000131 Length: 1653

Score: 131.00 Matches: 66

Percent Similarity: 33.8% Conservative: 41

Best Local Similarity: 20.8% Mismatches: 110

Query Match: 6.4% Indels: 100

DB: 3 Gaps: 12

US-10-780-703-2 (1-379) x US-09-248-796A-336 (1-1653)

Db 1291 GCTTCACGGATTTTA-----ACACACACATAACCCCAAGA 1329
QY 338 rGlnAlaLeu 342
Db 1330 TTTGGCGACAGTC 1342

RESULT 15
US-09-270-767-12547
; Sequence 12547, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12547
; LENGTH: 934
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-12547

Alignment Scores:
Pred. No.: 0.000221 Length: 934
Score: 125.50 Matches: 52
Percent Similarity: 42.7% Conservative: 15
Best Local Similarity: 33.1% Mismatches: 54
Query Match: 6.2% Indels: 36
DB: 3 Gaps: 7

US-10-780-703-2 (1-379) x US-09-270-767-12547 (1-934)

QY 219 AsnSerSerThrSerArgLeuAlaLeuArgGlnGlnHisSerSerSer 238
Db 8 AACGGAGACGAGCAGC-----AGCAGCAGCAGCAGC 40
QY 239 AsnHisSerAspAsnLeuAsnAsnAsnAsnAsnLeuLeuLeuSer 258
Db 41 AATAGTGGTAACACACACACACACACACATCAATCAAT----- 85
QY 259 ThrGluTyrSerGlyAspGlySerThrThrThrThrThrThrThrThr 278
Db 86 -----AACACTATACCAACAGCAGCAGCAGCAGCAGCAGCAGC 124
QY 279 ThrIleAlaLeuAlaAsnGlnAsnIleTyrArgProMetProTyrAspThrSerAsnAsn 298
Db 125 TCATCGGCC-----AAACCCCTACTTATCGCGACCCCTACAGCTCGCTAAATCGG 172
QY 299 ThrLeuIleValSerThrArgAsnHisGlnAspAspAsp-----GluThrAlaIleValAsp 317
Db 173 GTGCTC-----AATTCGAGCGGATTCCTGGAGATCCTGCGAGCAGCAA 217
QY 318 AspLeuGlnArgLeuValAsn-TyrGlnIleSerAspGlyAlaThrThrLeuMetProG1 337
Db 218 CAACACAGCAGCAGCTCAACGGGACCACTCAGCGCAACAGCAGCAGCAGCAGCAGCAGCAGC 277
QY 337 nThrGlnAlaLeuAlaMetAsnMetIlePro-----AlaGlyThrIleProAsnAs 355
Db 278 AGCGGGCGGACGACTAGCTCGCGAGATTCACAGCAGATCGGGACATCAATCACTACCC 337
QY 355 nAlaLeuTrpAspMetTrpAsnProIleValProAspGlyAsnArgAsp 371
Db 338 GCCAATA-----GCAGTTCAACCAACGCGACCCGAGCGAGGACGAGGAA 380

Search completed: April 6, 2006, 10:02:43
Job time : 249 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 6, 2006, 09:54:37 ; Search time 4958 Seconds

(without alignments)
3576.503 Million cell updates/sec

Title: US-10-780-703-2

Perfect score: 2034

Sequence: 1 MAIVSSTTSIIIPMSNQVNN.....MNPVIVDGNRHYTNIPFK 379

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q/abs/ABSWEB spool/US10780703/runat_06042006_101339_6361/app_query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs03h
-USER=US10780703 @CGN_1_1_5315 @runat_06042006_101339_6361 -NCFU=6 -ICFU=3
-NO_MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gse1:*
10: gb_gse2:*
11: gb_gse3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 1498 | 73.6 | 947 | 8 | DR749909 |
| 2 | 1374 | 67.6 | 901 | 8 | DR749910 |
| 3 | 1113 | 54.7 | 622 | 1 | AV827617 |
| 4 | 882.5 | 43.4 | 880 | 8 | DR938316 |
| 5 | 881.5 | 43.3 | 648 | 1 | AW683672 |
| 6 | 878.5 | 43.2 | 653 | 2 | BE204243 |
| 7 | 877 | 43.1 | 906 | 8 | DR935181 |

| | | | | | |
|----|-------|------|------|----|----------|
| 8 | 871 | 42.8 | 853 | 8 | DR928151 |
| 9 | 814.5 | 40.0 | 1404 | 10 | CL960332 |
| 10 | 802.5 | 39.5 | 623 | 8 | CX527114 |
| 11 | 799 | 39.3 | 516 | 3 | BI893633 |
| 12 | 797.5 | 39.2 | 1245 | 10 | CL971389 |
| 13 | 789.5 | 38.8 | 694 | 2 | BF625246 |
| 14 | 782 | 38.4 | 525 | 8 | DT006901 |
| 15 | 765 | 37.6 | 801 | 8 | DR813028 |
| 16 | 761 | 37.4 | 784 | 6 | CB621699 |
| 17 | 759.5 | 37.3 | 780 | 7 | CN149464 |
| 18 | 756.5 | 37.2 | 763 | 6 | CB649362 |
| 19 | 753.5 | 37.0 | 770 | 6 | CB621762 |
| 20 | 740.5 | 36.4 | 555 | 1 | AW443468 |
| 21 | 740.5 | 36.4 | 649 | 7 | CK077495 |
| 22 | 737 | 36.2 | 742 | 8 | DR787066 |
| 23 | 725.5 | 35.7 | 640 | 7 | CN144198 |
| 24 | 712.5 | 35.0 | 657 | 2 | BI268158 |
| 25 | 710 | 34.9 | 486 | 8 | DT004645 |
| 26 | 707.5 | 34.8 | 741 | 7 | CO096386 |
| 27 | 696.5 | 34.2 | 690 | 9 | BA481683 |
| 28 | 668.5 | 32.9 | 813 | 9 | AZ126200 |
| 29 | 664.5 | 32.7 | 494 | 7 | CK076588 |
| 30 | 657 | 32.3 | 427 | 1 | AV440745 |
| 31 | 657 | 32.3 | 468 | 1 | AV442690 |
| 32 | 644.5 | 31.7 | 823 | 9 | CC659605 |
| 33 | 631.5 | 31.0 | 633 | 2 | BG320780 |
| 34 | 631.5 | 31.0 | 879 | 9 | CC589779 |
| 35 | 627 | 30.8 | 838 | 9 | BH666975 |
| 36 | 618 | 30.4 | 879 | 9 | CC594458 |
| 37 | 615 | 30.2 | 853 | 9 | CC637855 |
| 38 | 615 | 30.2 | 929 | 9 | CC637864 |
| 39 | 614 | 30.2 | 901 | 10 | CG180706 |
| 40 | 612 | 30.1 | 973 | 9 | CC714199 |
| 41 | 611 | 30.0 | 923 | 10 | CG195779 |
| 42 | 605 | 29.7 | 859 | 10 | CG180704 |
| 43 | 600.5 | 29.5 | 739 | 10 | CL155695 |
| 44 | 594.5 | 29.2 | 540 | 8 | DN183182 |
| 45 | 591 | 29.1 | 986 | 9 | CC339647 |

ALIGNMENTS

DR749909 947 bp mRNA linear EST 19-JUL-2005
85-L020254-065-002-E11-SeLA MP12-ADIS-065d Arabidopsis thaliana
CDNA clone 002-E11, mRNA sequence.
DR749909
DR749909.1 GI:71035249
EST.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 947)
Paz-Ares, J., Valencia, A., Costantino, P., Vittorioso, P., Davies, B.,
Gilmartin, P., Giraudat, J., Parcy, F., Reindl, A., Sablowski, R.,
Cupland, G., Martin, C., Angenot, G.C., Baumelein, H., Mock, H.P.,
Carbonero, P., Colombo, L., Tonelli, C., Engstroem, P.,
Droge-Laser, W., Gatz, C., Kavanagh, T., Kuhnir, S., Zabeau, M.,
Lau, T., Hordsworth, M., Ruberti, I., Ratcliff, P., Smeekens, S.,
Somssich, I., Weishaar, B. and Traas, J.
REGIA, an EU project on functional genomics of transcription
factors from Arabidopsis thaliana
Comp. Funct. Genomics 3 (2), 102-108 (2002)
Contact:
Paz-Ares, Costantino, Vittorioso, Davies, Gilmartin, Giraudat, Parcy, Sabl
owski, Cupland, Martin, Angenot, Baumelein, Carbonero, Colombo, Tonelli,
Engstroem, Droge-Laser, Gatz, Kavanagh, Kuhnir, Zabeau, Lau, Hordsworth,
Ruberti, Smeekens, Somssich, Weishaar, Traas
Bielefeld University, Institute for Genome Research
Universitaetsstrasse 25, D-33594 Bielefeld, Germany

Email: bernd.weishaar@uni-bielefeld.de
 AGI: AtfG02450; Seqanalysis: undetermined frame; Translation: no full cds detected
 Data analysis performed in the frame of REGULATORS (Exploiting inter-species conservation in promoter sequences to identify regulators of reproductive development and physiological performance), a Triilateral Co-Operation in Plant Genomics between Spain (MCYT), France (GENOPLANTE) and Germany (GABI) coordinated by G. Coupland (coupland-ad-mpiz-koeln.mpg.de). Authors: Vincent Thureau (IBP-Orsay UMR8618 CNRS-UPS, thureau-ad-ibp.u-psud.fr) and Alain Lecharny (URGV-Evry UMR INRA-CNRS-UEVE, lecharny-ad-ibp.u-psud.fr).

Definition of the terms used to describe the quality of the clone: The about 2250 sequences from the clone collection were sorted according to clones and clustered. If more than one contig was formed, the clone was designated 'contamination'. The contigs and singletons were blasted against CDS plus pseudogenes from the TIGRv5 annotation, and the resulting AGI code is presented if more than 90 percent identity was found. The sequences were also blasted against all TIGRv5 introns, and matches longer than 50 bp with 95 percent identity are reported as 'intron found'. The remaining terms for Seqanalysis describe the outcome of the evaluation of the CDS detected after pairwise alignment with CDS plus pseudogenes from the TIGRv5 annotation file. The sequences or contigs for which a full CDS with or without STOP codon was detected, a BLASTp against all TIGRv5 protein sequences was performed. Full perfect: 100 percent identity; full good: better than 95 percent identity over more than 95 percent of the sequence; partial good: better than 95 percent identity over less than 95 percent of the sequence; weak similarity: less than 95 percent identity over less than 95 percent of the sequence; no similarity: no hit from BLASTp. Note that the collection contains a few clones for which sequencing was not successful, which could have simple technical reasons. Obviously, information about these clones is missing in the submitted data.

Insert Length: 947 Std Error: 0.00

Seq primer: SELA TCGCGTTAAGCTAGCATGATCTC.

FEATURES

Location/Qualifiers
 1..947
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /ecotype="Columbia"
 /db_xref="GABI:1452250"
 /db_xref="taxon:3702"
 /clone="002-E11"
 /lab_host="E. coli DH5alpha"
 /clone_lib="MP1Z-ADIS-065d"
 /note="Vector: pDONR201; In the context of the EU-funded project REGIA (QLG-C11999-00876, coordinator Javier Paz-Ares), a set of transcription factor ORFs was generated. The ORFs were produced in a decentralized way in the labs of the participants. Most of the ORFs were generated by RT-PCR using cDNA from various A. thaliana tissues as a template. Initially, it was planned to use yeast recombination to move the ORFs from the cloning vectors into target constructs. For this reason, a number of the clones contain 'REGIA tags' (RG tag):
 Spr-AATTCAGCTCACCAC-3pr; RG tag2:
 Spr-CATCGCAATCCGGGATC-3pr). During the lifetime of the project, the GATEWAY system became available and finally all ORFs were transferred into GATEWAY vectors. At the end of the REGIA project, the clones were collected as plasmid DNA. For end-sequencing at the DNA core facility of the MPI of Plant Breeding Research (ADIS; head: Bernd Weishaar), the plasmids were re-transformed into DH5alpha, arrayed into 96-well plates and new plasmid DNA was prepared. Re-transformation proved to be difficult because a number of clones displayed poor growth. Data submission has been handled by Gabipd (http://gabi.rzpd.de)."

ORIGIN

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|-----|
| Pred. No.: | 2,77e-146 | Length: | 947 |
| Score: | 1498.00 | Matches: | 277 |
| Percent Similarity: | 99.6% | Conservative: | 3 |
| Best Local Similarity: | 98.6% | Mismatches: | 1 |
| Query Match: | 73.6% | Indels: | 0 |
| DB: | 8 | Gaps: | 0 |

US-10-780-703-2 (1-379) x DR749909 (1-947)

| | | | |
|----|-----|--|-----|
| Qy | 13 | MetSerAsnGlnValAsnAsnAsnGluLysGlyIleGluAspAsnAspHisArgGlyGly | 32 |
| Db | 104 | ATGAGTAACCAAGTCAACAATAACGAAAAAGGTATAGAACATATGATCATAGAGCGGC | 163 |
| Qy | 33 | GlnGluSerHisValGlnAsnGluAspGluAlaAspHisAspHisAspMetValMet | 52 |
| Db | 164 | CAGAGAGTCAATGTCCAAATATGAAGTGAAGTCAATGATCATGATCATGATGATGATG | 223 |
| Qy | 53 | ProGlyPheArgPheHisProThrGluGluLeuLeuGluPheThrLeuArgArgLys | 72 |
| Db | 224 | CCCGAATTTAGATTCCATCTCTACCGAAGAAGAACTCATAGAGTTTACCTTCGCGCAAAA | 283 |
| Qy | 73 | ValGluGlyLysArgPheAsnValGluLeuLeuThrPheLeuAspLeuThrArgTyrAsp | 92 |
| Db | 284 | GTTGAGGCAACGCTTTAATGTAGAACTCATCTCTCTCGATCTTTATCGCTATGAT | 343 |
| Qy | 93 | ProThrGluLeuProAlaMetAlaAlaIleGlyGluLysGluThrPheThrValPro | 112 |
| Db | 344 | CTTTGGGAATCTCTCTGATATGGCGCGATAGGAGAGAAAGTGTACTTCTATGTGCCA | 403 |
| Qy | 113 | ArgAspArgLysThrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrp | 132 |
| Db | 404 | AGAGATCGGAATATAGAAATGGAGATAGACCGAAGCAGTAACGACTTCAGGATATTGG | 463 |
| Qy | 133 | LysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLys | 152 |
| Db | 464 | AAAGCCACCGAGCTGATAGATGATCAGATCGGAGACTTCTCGCGCTATCGGATTAAG | 523 |
| Qy | 153 | LysThrLeuValPheThrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMet | 172 |
| Db | 524 | AAAACCTTAGTTTCTACTCTGGTAAAGCCCTTAAGGCACTCTGACTAGTTGGATCATG | 583 |
| Qy | 173 | AsnGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLysAlaGluIleSerLeu | 192 |
| Db | 584 | AACGAGTATCGTCTCTCGTACCATGAACCGAGAAGTACCAAAAGGCTGAAATATCATTG | 643 |
| Qy | 193 | CysArgValTyrLysArgProGlyValGluAspHisProSerValProArgSerLeuSer | 212 |
| Db | 644 | TGCCGAGTGTACAAAGCCGAGGAGTAGAAGATCATCATCGGTACCACGTTCTCTCTCC | 703 |
| Qy | 213 | ThrArgHisAsnHisAsnSerThrSerSerArgLeuAlaLeuArgGlnGlnGln | 232 |
| Db | 704 | ACAAGATCATATACCACTAATCTCATCGACATCATCCGTTTAGCTTAAGACACACACAA | 763 |
| Qy | 233 | HisHisSerSerSerSerAsnHisSerAspAsnAsnLeuAsnAsnAsnAsnIleAsn | 252 |
| Db | 764 | CACCATTCATCTCTCTAATCATTTCCGACAAACACCTTAACAAACAACAACATTCAC | 823 |
| Qy | 253 | AsnLeuGluLysLeuSerThrGluTyrSerGlyAspGlySerThrThrThrThrThr | 272 |
| Db | 824 | AATCTCAGAGAGCTCTCCACCGAATATTCGCGCCACCGCAGCAACAAACACACACCA | 883 |
| Qy | 273 | AsnSerAsnSerAspValThrIleAlaLeuAlaAsnGlnAsnIleTyrArgProMetPro | 292 |
| Db | 884 | AACAGTAATCTGACGTTACCATTTGCTTAGCCCATCAAAACATATATATCGTCCATGCT | 943 |
| Qy | 293 | Tyr | 293 |
| Db | 944 | TAC | 946 |

RESULT 2

DR749910/c

LOCUS

DEFINITION 85-L020255-065-002-E11-SeLb MP1Z-ADIS-065d Arabidopsis thaliana

DR749910

901 bp mRNA linear

EST 19-JUL-2005

```

cdna clone 002-E11, mRNA sequence.
DR749910
DR749910.1 GI:71035250
EST
SOURCE
Arabidopsis thaliana (chale cress)
Arabidopsis thaliana
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 901)
Paz-Ares,J., Valencia,A., Costantino,P., Vittorioso,P., Davies,B.,
Gilmartin,P., Giraudat,J., Parcy,F., Reindl,A., Sablowski,R.,
Coupland,G., Martin,C., Angenent,G.C., Baumelein,H., Mock,H.P.,
Carbonero,P., Colombo,L., Tonelli,C., Engstrom,P.,
Dröge-Laser,W., Gatz,C., Kavanagh,T., Kushnir,S., Zabeau,M.,
Laux,T., Hordeworth,M., Ruberti,I., Ratcliff,F., Smeekens,S.,
Sonsiech,I., Weisshaar,B. and Traas,J.
REGIA, an EU project on functional genomics of transcription
factors from Arabidopsis thaliana
Comp. Funct. Genomics 3 (2), 102-108 (2002)
Contact:
Paz-Ares, Costantino, Vittorioso, Davies, Gilmartin, Giraudat, Parcy, Sablowski, Coupland, Martin, Angenent, Baumelein, Carbonero, Colombo, Tonelli, Engstrom, Dröge-Laser, Gatz, Kavanagh, Kushnir, Zabeau, Laux, Holdsworth, Ruberti, Smeekens, Somsich, Weisshaar, Traas
Bielefeld University, Institute for Genome Research
Universitätsstrasse 25, D-33594 Bielefeld, Germany
Email: bernd.weisshaar@uni-bielefeld.de
AGI: AT2002450; SeqAnalysis: undetermined frame; Translation: no full cds detected
Data analysis performed in the frame of REGULATORS (Exploiting inter-species conservation in promoter sequences to identify regulators of reproductive development and physiological performance), a Trilateral Co-Operation in Plant Genomics between Spain (MCYT), France (GENOPLANTE) and Germany (GABI) coordinated by G. Coupland (coupland-ad-mpiz-koeln.mpg.de). Authors: Vincent Thareau (IBP-Orsay UMR8618 CNRS-UPS, thareau-ad-ibp.u-psud.fr) and Alain Lecharny (URGV-Evry UMR INRA-CNRS-UEVE, lecharny-ad-ibp.u-psud.fr).
Definition of the terms used to describe the quality of the clones:
The about 2250 sequences from the clone collection were sorted according to clones and clustered. If more than one contig was formed, the clone was designated 'Contamination'. The contigs and singletons were blasted against CDS plus pseudogenes from the TIGRv5 annotation, and the resulting AGI code is presented if more than 90 percent identity was found. The sequences were also blasted against all TIGRv5 introns, and matches longer than 50 bp with 95 percent identity are reported as 'intron found'. The remaining terms for SeqAnalysis describe the outcome of the evaluation of the CDS detected after pairwise alignment with CDS plus pseudogenes from the TIGRv5 annotation file. The sequences or contigs for which a full CDS with or without STOP codon was detected, a BLASTp against all TIGRv5 protein sequences was performed. Full perfect: 100 percent identity; full good: better than 95 percent identity over more than 95 percent of the sequence; partial good: better than 95 percent identity over less than 95 percent of the sequence; weak similarity: less than 95 percent identity over less than 95 percent of the sequence; no similarity: no hit from BLASTp. Note that the collection contains a few clones for which sequencing was not successful, which could have simple technical reasons. Obviously, information about these clones is missing in the submitted data.
Insert Length: 901 Std Error: 0.00
Seq primer: SeUB GTAACATCAGAGATTGTGAGACAC.
Location/Qualifiers
1. 901
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="GABI:1452251"
/db_xref="taxon:3702"
/clones="002-E11"
/lab_host="E. coli DH5alpha"

```

/clone lib="MPIZ-ADIS-065d"
/notes=Vector: pDONR201; In the context of the EU-funded project REGIA (QLG-CT11999-00876, coordinator Javier Paz-Ares), a set of transcription factor ORFs was generated. The ORFs were produced in a decentralized way in the labs of the participants. Most of the ORFs were generated by RT-PCR using cDNA from various *A. thaliana* tissues as a template. Initially, it was planned to use yeast recombination to move the ORFs from the cloning vectors into target constructs. For this reason, a number of the clones contain 'REGIA tags' (RG_tag):
Spr-AATCGAGTCCACACC-3pr; RG_tag2:
Spr-CATGCGCAATCCGGGATC-3pr). During the lifetime of the project, the GATEWAY system became available and finally all ORFs were transferred into GATEWAY vectors. At the end of the REGIA project, the clones were collected as plasmid DNA. For end-sequencing at the DNA core facility of the MPI of Plant Breeding Research (ADIS; head: Bernd Weisshaar), the plasmids were re-transformed into DH5alpha, arrayed into 96-well plates and new plasmid DNA was prepared. Re-transformation proved to be difficult because a number of clones displayed poor growth. Data submission has been handled by GabiPD (<http://gabi.rzpd.de>).

ORIGIN

Alignment Scores:
Pred. No.: 2,66e-133 Length: 901
Score: 1374.00 Matches: 259
Percent Similarity: 98.9% Conservative: 2
Best Local Similarity: 98.1% Mismatches: 3
Query Match: 67.6% Indels: 0
DB: 8 Gaps: 0

US-10-780-703-2 (1-379) x DR749910 (1-901)

QY 113 ArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrp 132
DB 901 AGAGATCGGAATATAGAAATGAGATAGACCGACCGAGTACCGACTTCAGGATATTGG 842
QY 133 LysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLys 152
DB 841 AAAGCCACCGAGCTGATGGATGATCAGATCGGAGACTTCTCGGCCCTATCGGATTAAAG 782
QY 153 LysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMet 172
DB 781 AAAACCCCTAGTTTCTACTCTGGTAAGCCCCCTAAAGGCACCTCGTACTAGTTGGATCATG 722
QY 173 AsnGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLysAlaGluIleSerLeu 192
DB 721 AACGAGTATCGTTCCTCGTACCATGAACCGAGAGTACCACCAAGGCTGAAATATCATTTG 662
QY 193 CysArgValTyrLysArgProGlyValGluAspHisProSerValProArgSerLeuSer 212
DB 661 TGCCGAGTGTTACAAAAGGCCAGGAGTAGAAGATCATCCATCGGTACCCAGCTTCTCTCTCC 602
QY 213 ThrArgHisHisAsnHisAsnSerSerThrSerSerArgLeuAlaLeuArgGlnGlnGln 232
DB 601 ACAAGACATCATCAACATACTCATCGACATCATCCGTTTACCGTTTAAAGCAACAACAA 542
QY 233 HisHisSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 252
DB 541 CACCATTCATCT 482
QY 253 AsnLeuGluLysLeuSerThrGluTyrSerGlyAspGlySerThrThrThrThrThrThr 272
DB 481 AATCTCGAGAAGCTCTCCACCGAATATTCCGGCGAGCGGAGCAGCAACAACAAGCACCACA 422
QY 273 AsnSerAsnSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 292
DB 421 AACAGTAACTCTGACGCTTACCATTTGCTTAGCCCAATCAAAACATATATATCGTCAATCCCT 362
QY 293 TyrAspThrSerSerAsnAsnThrLeuIleValSerThrArgAsnHisGlnAspAspGlu 312

```

|||||
361 TACGACACAGCAACACACATTTAGTCTCTACGAGAAATCATCAAGACGATGAA 302
QY 313 ThrAlaIleValAspAspLeuGlnArgLeuValAsnTyrGlnIleSerAspGlyAlaThr 332
Db 301 ACTGCCATTTGTTGACGATCTTCAAGACATAGTTAACTACCAATATCAGATCAGCGACA 242
QY 333 ThrLeuMetProGlnThrGlnAlaLeuAlaMetAsnMetIleProAlaGlyThrIle 352
Db 241 AGCTTAATGCTCAAACTCAAGCGGCTTAGCTATGAACATGATCTCTGCAGGAACGAT 182
QY 353 ProAsnAsnAlaLeuTrpAspMetTrpAsnProIleValProAspGlyAsnArgAspHis 972
Db 181 CCAACAATGCTTTGGGGGATATGTGGATCCCAATAGTACCAGATGGAACACAGATCAC 122
QY 373 TyrThrAsnIle 376
Db 121 TATACATAACATG 110

RESULT 3
AV827617 622 bp mRNA linear EST 01-APR-2002
LOCUS AV827617 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-17-121 5',
DEFINITION mRNA sequence.
ACCESSION AV827617.1 GI:19869677
VERSION AV827617
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 622)
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.,
and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@tc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda PLC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
details.

FEATURES
source Location/Qualifiers
1..622
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL09-17-121"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="PH108"
/clone_lib="RAFL9"
/note="Site 1: BamHI; Site 2: SalI; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"

ORIGIN
Alignment Scores: 3.89e-106 Length: 622
Pred. No.: 1113.00 Matches: 206
Score: 100.0% Conservative: 0
Percent Similarity: 100.0%
Best Local Similarity: 100.0% Mismatches: 0

```

```

Query Match: 54.7% Indels: 0
DB: 1 Gaps: 0

US-10-780-703-2 (1-379) x AV827617 (1-622)

QY 88 LeuTyrArgTyrAspProTrpGluLeuProAlaMetAlaIleGlyGluLysGluTrp 107
Db 3 CTTTATCGCTATGATCTCTTGGAACTTCTCTGCTATGGCGCGATAGAGAGAGAGAGTGG 62
QY 108 TyrPheTyrValProAspAspArgLysTyrArgAsnGlyAspArgProAsnArgValThr 127
Db 63 TACTTTCATGTGCGCAAGAGATCGGAATATAGAATGGATATAGACCGAACCGATACG 122
QY 128 ThrSerGlyTyrTrpLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArg 147
Db 123 ACTTCAGGATATGGAAGGCCACCGAGCTCATAGGATGATCAGATCGAGACTTCTCGG 182
QY 148 ProIleGlyLeuLysLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArg 167
Db 183 CCTATCGGATTAAGAAACCCCTAGTTTCTTCTGTTAAAGCCCTTAAAGGCACTCGT 242
QY 168 ThrSerTrpIleMetAsnGluTyrArgLeuProHisGluThrGluLysTyrGlnLys 187
Db 243 ACTAGTTGGATCATGAACGATATCGTCTTCCGACCATGAACCGAGAGTACCCAAAG 302
QY 188 AlaGluIleSerLeuCysArgValTyrLysArgProGlyValGluAspHisProSerVal 207
Db 303 GCTGAATATCATTTGTCCGAGTGACAAAGGCCAGGAGTAGAAGATCATCCATCGGTA 362
QY 208 ProArgSerLeuSerThrArgHisHisHisAsnSerSerThrSerSerArgLeuAla 227
Db 363 CCAGCTTCTCTCTCCACAAACATCAATCACTCATCGACATCATCCCGTTTATGCC 422
QY 228 LeuArgGlnGlnHisHisSerSerSerSerSerSerSerSerSerSerSerSerSerSer 247
Db 423 TTAAGACAAACAAACCAACCATTCATCTCTCTTATCATTCATTCGACACACACCTTAA 482
QY 248 AsnAsnAsnIleAsnAsnLeuGluLysLeuSerThrGluTyrSerGlyAspGlySerThr 267
Db 483 AACAAACACATCAACATCTCGAGAGCTCTCCACCGAATATTCGGCGGACGAGCACA 542
QY 268 ThrThrThrThrThrAsnSerAsnSerAspValThrIleAlaLeuAlaAsnGlnAsnIle 287
Db 543 ACAACACACGACACAAACAGTAACTCTGACGTTACCATTTGCTCTAGCCCAATCAAA 602
QY 288 TyrArgProMetProTyr 293
Db 603 TATCGTCCAATGCGCTTAC 620

RESULT 4
DR938316 880 bp mRNA linear EST 02-AUG-2005
LOCUS DR938316 Aquilegia cDNA library Aquilegia formosa x Aquilegia
DEFINITION pubescens cDNA clone COLPM65, mRNA sequence.
ACCESSION DR938316.1 GI:71707679
VERSION DR938316
KEYWORDS EST.
SOURCE Aquilegia formosa x Aquilegia pubescens
ORGANISM Aquilegia formosa x Aquilegia pubescens
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
Ranunculaceae; Aquilegia.
1 (bases 1 to 880)
Hodges, S.A., Rensink, W., Buell, C.R., Borevitz, J., Kramer, E.,
Nordborg, M., and Tomkins, J.
Generation of ESTs from Aquilegia
Unpublished (2005)
Other ESTs: EST1129854
Contact: Scott Hodges
Department of Ecology, Evolution and Marine Biology
University of California, Santa Barbara
Santa Barbara, CA 93106, USA
Tel: 805 893 7813

```

| | | | | |
|------------|-----|---|-----|--|
| Db | 427 | CCTCAGCAAGACACTGATAGATACCAAAAGGCGAGAAATTCGCTTTGTTCGAGTCTACAG | 486 | |
| QY | 198 | ArgProGlyValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsn | 217 | |
| Db | 487 | AGACCTGGAGTAGAAGACCAAGCCAGAGTCCCGGCTCC | 528 | |
| QY | 218 | HisAsnSerSerThrSerSerArgLeuAlaLeuArgGlnGlnHisHisSerSerSer | 237 | |
| Db | 529 | CATCTCTTCGACCATCT | 555 | |
| QY | 238 | SerAsnHisSerAspAsnAsnLeuAsnAsn | 255 | |
| Db | 556 | GGAACCTCAGCCGGATAAAGACAGCAGAACTCATCTATGGAATTTCAAGCTTTCAA | 615 | |
| QY | 256 | LysLeuSerThrGluTyrSerGlyAspGlySerThrThrThrThrThrThrThrThr | 275 | |
| Db | 616 | GCATAGATAGCAACAACTCACCCGCACTACTACTACTACTACTACTACTACTACT | 675 | |
| QY | 276 | SerAspValThrIleAlaLeuAlaAsn | 293 | |
| Db | 676 | GAAAAGCTGAATGAACGACGATGCAAGCAGTACTAGCAATGA-TACAGGAACCTGCTCTGG | 734 | |
| QY | 294 | AspThr | 308 | |
| Db | 735 | GATAGCAATGTTCCTCCATGATCATGATTCGCGCAGTATTGTTCACAGCTTCAACCATTTGAG | 794 | |
| QY | 309 | AspAspAspGluThrAlaIleValAspAspLeuGlnArgLeuValAsnTyrGlnIleSer | 328 | |
| Db | 795 | GAGGTCAGTGGTACTGCTCATTCATTCAACAATATTCAAGG | 833 | |
| QY | 329 | AspGlyAlaThrThrLeuMetProGlnThrGlnAlaLeuAlaMetAsnMetIlePro | 348 | |
| Db | 834 | | 851 | |
| QY | 349 | AlaGlyThrIleProAsnAsnAlaLeu | 357 | |
| Db | 852 | TCAAATACTTGTAAATAACAATAACCCCTC | 878 | |
| RESULT 5 | | | | |
| LOCUS | | | | |
| DEFINITION | | | | |
| ACCESSION | | | | |
| VERSION | | | | |
| KEYWORDS | | | | |
| SOURCE | | | | |
| ORGANISM | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| COMMENT | | | | |
| FEATURES | | | | |
| SOURCE | | | | |

Fax: 805 893 4724
Email: hodge@lifesci.ucsb.edu
Seq primer: M13 Reverse
Location/Qualifiers
1. .880
/organism="Aquilegia formosa x Aquilegia pubescens"
/mol_type="mRNA"
/db_xref="taxon:338618"
/clone="CO1PM65"
/tissue_type="mixed shoot and floral apical meristems, flower buds, leaves and roots"
/lab_host="DH10B T1 (T1 and T5 phage resistance)"
/clone_lib="Aquilegia cDNA library"
/notes="Vector: pCMV SPOR6.1; Site 1: EcoRI; Site 2: NotI; F2, F3, and F4 lines of Aquilegia formosa X A. pubescens were grown from seed in greenhouses at UC Santa Barbara. From these plants three sets of tissue were collected: 1) Small flower buds (<10 mm) and very young inflorescences (71 & 29% by weight respectively), 2) Medium (7-20 mm) and large (at or near anthesis) flower buds (65 & 35% by weight respectively) and 3) Shoot apical meristems. A fourth set of tissue was collected from plants of A. formosa. These plants were grown from seed in sand and at approximately 1 month root tissue and leaf tissue of various developmental stages were collected (84 & 16% by weight respectively). Total RNA was extracted from each set of tissue and pooled in the following proportions: 1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled total RNA, mRNA was extracted and enriched for full-length messages and then normalized with proprietary methods by Invitrogen."

| | | | | |
|--|----------|--|-----|--|
| Alignment Scores: | | | | |
| Pred. No.: | 9.92e-82 | Length: | 880 | |
| Score: | 882.50 | Matches: | 186 | |
| Percent Similarity: | 65.3% | Conservative: | 29 | |
| Best Local Similarity: | 56.5% | Mismatches: | 67 | |
| Query Match: | 43.4% | Indels: | 48 | |
| DB: | 8 | Gaps: | 7 | |
| US-10-780-703-2 (1-379) x DR938316 (1-880) | | | | |
| QY | 38 | GlnAsnGluAspGluAlaAspAspHisAspMetValMetProGlyPheArgPhe | 57 | |
| Db | 7 | CAAAACACGACGACAGGATACACGACGACGATGCTGCTGCTTCGGTTT | 66 | |
| QY | 58 | HisProThrGluGluLeuLeuGluPheTyrLeuArgLysValGluGlyLysArg | 77 | |
| Db | 67 | CACCAACCGAGGAGAGAACTCATTGAATCTATCTCCCGCTAAGGTTGAGGCGCAAGCGC | 126 | |
| QY | 78 | PheAsnValGluLeuIleThrPheLeuAspLeuTyrArgTyrAspProThrGluLeuPro | 97 | |
| Db | 127 | TTTAATGTGGAGCTTATTACTTCTCGATCTATATCGCTATGACCCCTGGAGCTTCCA | 186 | |
| QY | 98 | AlaMetAlaAlaIleGlyGlyLysGluTyrPheTyrValProArgAspArgLysTyr | 117 | |
| Db | 187 | GCATTAGTCCCATGAGAGAAAGATGTTCTTCTATGTCCAGAGACCGAAAGTAT | 246 | |
| QY | 118 | ArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrpLysAlaThrGlyAla | 137 | |
| Db | 247 | CGCAACGGAGATCGCCCTAACCGGGTGACACATCTGGCTACTGGAAGGCTACTGTGCA | 306 | |
| QY | 138 | AspArgMetIleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPhe | 157 | |
| Db | 307 | GACAGAATGATCCGAAATGATTAACCTTCAGATCCATTCGGCTTAAGAAGACACTAGTCTTC | 366 | |
| QY | 158 | TyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeu | 177 | |
| Db | 367 | TATTCGGGGAAGACCTTAAGGCATCAGAACAGTTGGATCATGATGATGATGATGAT | 426 | |
| QY | 178 | ProHisGluThrGluLysTyrGlnLysAlaGluLeuSerLeuCysArgValTyrLys | 197 | |


```

|||||
255 TATCGTATGACCTTGGAACTTCTGCTTTGGCAGCAATCGGAGAGAGAGAGTGTGAC 314
|||
109 PheTyrValProArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThr 128
|||
315 TTTTATGTCCTCGAGATAGGAAGTATCGAAACGGTGTATCGTCCAAATCGTGTAACT 374
|||
129 SerGlyTyrTyrLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgPro 148
|||
375 TCTGGTATTGGAAGGCAACAGGCGGTAGGATTCGAACTGAAAACCTTCGGTCA 434
|||
149 IleGlyLeuLysLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThr 168
|||
435 ATTGACTCAAGAAACCTAGTTCTTCTTCTGGAAGGCTCTTAAAGGCATCGAAC 494
|||
169 SerTrpIleMetAsnGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLysAla 188
|||
495 AGTTGGATTATGAATGAGTATAGATTATCCCAACATGAACCGAATGATCAAAAGGGT 554
|||
189 GluIleSerLeuCyAspValTyrLysArgProGlyValGluAspHisProSerValPro 208
|||
555 GAGATATCCCTATGTCGTGTATACAGAGAGCAGGATAGAGATCATCTCCCTCCCT 614
|||
209 ArgSerLeuSer---ThrArg-HisHisAsnHisAsn 219
|||
615 CGTGTCTCCCAATTACAGCACCATCATCATCAT 651
|||

RESULT 7
DR935181
LOCUS
DEFINITION
EST1126720 Aquilegia cDNA library Aquilegia formosa x Aquilegia
pubescens cDNA clone COLP234, mRNA sequence.
DR935181
ACCESSION
DR935181.1 GI:71704544
VERSION
EST.
KEYWORDS
SOURCE
ORGANISM
Aquilegia formosa x Aquilegia pubescens
Aquilegia formosa x Aquilegia pubescens
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
Ranunculaceae; Aquilegia.
REFERENCE
1 (bases 1 to 906)
Hodges,S.A., Rensink,W., Buell,C.R., Borevitz,J., Kramer,E.,
Nordborg,M. and Tomkins,J.
Generation of ESTs from Aquilegia
Unpublished (2005)
Other ESTs: EST1126719
Contact: Scott Hodges
Department of Ecology, Evolution and Marine Biology
University of California, Santa Barbara
Santa Barbara, CA 93106, USA
Tel: 805 893 7813
Fax: 805 893 4724
Email: hodges@lifesci.ucsb.edu
Seq primer: M13 Reverse.
FEATURES
Location/Qualifiers
1..906
/organism="Aquilegia formosa x Aquilegia pubescens"
/mol_type="mRNA"
/db_xref="taxon:338618"
/clone="COLP234"
/tissue_type="mixed shoot and floral apical meristems,
flower buds, leaves and roots"
/lab_hosts="DH10B T1 (T1 and T5 phage resistance)"
/clone_lib="Aquilegia cDNA library"
/note="Vector: pCMV SPORT6.1; Site 1: EcoRI; Site 2: NotI;
F2, F3, and F4 lines of Aquilegia formosa x A. pubescens
were grown from seed in greenhouses at UC Santa Barbara.
Small flower buds (<10 mm) and very young inflorescences
(71 & 29% by weight respectively), 2) Medium (7-20 mm) and
large (at or near anthesis) flower buds (65 & 35% by
weight respectively) and 3) Shoot apical meristems, a
fourth set of tissue was collected from plants of A.

```

formosa. These plants were grown from seed in sand and at approximately 1 month root tissue and leaf tissue of various developmental stages were collected (84 & 16% by weight respectively). Total RNA was extracted from each set of tissue and pooled in the following proportions: 1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled total RNA, mRNA was extracted and enriched for full-length messages and then normalized with proprietary methods by Invitrogen."

ORIGIN

```

Alignment Scores:
Pred. No.: 3,91e-81 Length: 906
Score: 877.00 Matches: 174
Percent Similarity: 74.3% Conservative: 25
Best Local Similarity: 64.9% Mismatches: 35
Query Match: 43.1% Indels: 34
DB: 8 Gaps: 5

US-10-780-703-2 (1-379) x DR935181 (1-906)

Qy 38 GlnAsnGluAspGluAlaAspAspHisAspHisMetValMetProGlyPheArgPhe 57
|||
Db 143 CAAAACAACGACGACGAGGATAACCCAGCACGATATGTCATGCTGGCTTCGGTTT 202
|||
Qy 58 HisProThrGluGluGluLeuGluPheTyrIleuArgGlyValGluGlyArg 77
|||
Db 203 CACCAACCCGAGGAGAACTCATTTCTCCCGCTAAGGTTGAGGGCAAGCGC 262
|||
Qy 78 PheAsnValGluLeuIleThrPheLeuAspLeuTyrArgTyrAspProTrpGluLeuPro 97
|||
Db 263 TTTAATGTGGAGCTTATTACTTTCTCGATATATCGTATGACCTTGGAGCTTCCA 322
|||
Qy 98 AlaMetAlaIleGlyGluLysGluTyrPheTyrValProArgAspArgLysTyr 117
|||
Db 323 GCATTAGCTGCCATTGGAGAGAAATGGTTCTTCTATGTGCCAGCAGACCAATAT 382
|||
Qy 118 ArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrpLysAlaThrGlyAla 137
|||
Db 383 CGTAACGAGATGCCCTAACCGGTGCAACATCTGGCTACTGAGAGGCTACTGGTGA 442
|||
Qy 138 AspArgMetIleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPhe 157
|||
Db 443 GACAGATGATCCGAATGATACTTGAGATCCATTGGCTTAAAGAGACACTAGTCTTC 502
|||
Qy 158 TyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeu 177
|||
Db 503 TATTCCGGAAAGCACCTAAAGGCATCAGAACCAAGTTGGATCATGAATGATACCGTTTA 562
|||
Qy 178 ProHisHisGluThrGluLysTyrGlnLysAlaGluIleSerLeuCyAspValTyrLys 197
|||
Db 563 CCTCAGCAAGACACTGATAGTACCAAAAGCGAAATTTCCGCTTTTCGAGTCTACAG 622
|||
Qy 198 ArgProGlyValGluAspHisProSerValProArgSerLeuSerThrArgHisAsn 217
|||
Db 623 AGACCTGGAGTAGAGACCCAGCCAGAGTTCCCGGCTCC-----TCT 664
|||
Qy 218 HisAsnSerSerThrSerArgLeuAlaLeu-----ArgGlnGlnGlnHisHis 234
|||
Db 665 CATTTCTTCGACCATCTTCATCAAGAGGAACTCAGCCGGATTAAGACAGACGAACTGT 724
|||
Qy 235 SerSerSerSerAsnHisSerAspAsnAsnLeuAsnAsnAsnAsnAsnAsnLeu 254
|||
Db 725 ATCATTATGGAATTCAGGCTTTCAAGCATACATAGCAACCACTCA----- 772
|||
Qy 255 GluLysLeuSerThrGluTyrSerGlyAspGlySerThrThrThrThrThrThrTh 272
|||
Db 773 -----CCCGCAACTACTACTACTACTACTACTACTACTACTACTACTAGATG 805
|||
Qy 273 -----AsnSerAsnSerAspValThrIleAlaLeuAla 283
|||
Db 806 GAAAAGCTGAATGAGCAGATGCAAGCAGTACTAGCAATGATATACCGGAACCTCTCTGGG 865
|||

```

| | | | |
|------------|---|--|-----|
| 155 | DB | CAAAACACGACGACGAGGATAACACGACGACGATATGCTCATGCTCGCTTCGCTTC | 21 |
| 58 | QY | HisProThrGluGluGluLeuLeuGluPheTyrlieuArgArglyseValGluGlyLysArg | 77 |
| 215 | DB | CACCAACCCGAGGAGAACTCATTTGATTCCTCGCGTAAAGTTGAGGGCAAGCGC | 274 |
| 78 | QY | PhenValGluLeuLeuThrPheLeuAspLeuTyrlArgTyrlAspProTrpGluLeuPro | 97 |
| 275 | DB | TTTAAATGGAGACTTATTTCTTCTCGATCTATCGCTATGACCCCTGGGAGCTTCCA | 334 |
| 98 | QY | AlaMetAlaAlaIleGlyGluLysGluTyrlPheTyrlValProArgAspArgLysTyrl | 117 |
| 335 | DB | GCATTAGCTGCCTTGGAGAGAAAGATGGTTCTTCTATGTCGACGAGACCGAAATAT | 394 |
| 118 | QY | ArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrlTrpLysAlaThrGlyAla | 137 |
| 395 | DB | CGTAACGAGATCGCCCTTACCGGGGTGCAACATCTGGCTACTGGAAGGCTACTGGTGCA | 454 |
| 138 | QY | AspArgMetIleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPhe | 157 |
| 455 | DB | GACAGATGATCCGAAATGATAAATTCGAGATCCATTTGGGCTTAAAGAGACACTAGTCTTC | 514 |
| 158 | QY | TyrlSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrlArgLeu | 177 |
| 515 | DB | TATTCGGGAAGACACCTTAAGGATCAGAACCGATTTGGATCATGATGATGATCCGTTTA | 574 |
| 178 | QY | ProHisHisGluThrGluLysTyrlGlnLysAlaGluLysSerLeuCysArgValTyrlLys | 197 |
| 575 | DB | CCTCAGCAAGACACTGATAGATACCAAAAGGAGAAATTTGCTTGTTCGAGTCTACAAG | 634 |
| 198 | QY | ArgProGlyValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsn | 217 |
| 635 | DB | AGACCTGGAGTAGAAGACACCGCCAGAGATTTCCCGGCTCC-----TCT | 676 |
| 218 | QY | HisAsnSerSerThrSerArgLeuAlaLeu-----ArgGlnGlnHisHis | 234 |
| 677 | DB | CATTCTTTCGACCACTTCATCAAGGAACTCAGCCGGATAAAGACAGACGAACTCTG | 736 |
| 235 | QY | SerSerSerSerAsnHisSerAspAsnLeuAsnAsnLeuAsnLeuAsnLeu | 254 |
| 737 | DB | ATCACTTATGGAATTTCAAGCTTTCAAGCATACGATAGCAACAACTCA----- | 784 |
| 255 | QY | GluLysLeuSerThrGluTyrlSerGlyAspGlySerThrThrThrThrThrThr | 272 |
| 785 | DB | -----CCCGCAACTACTACTACTACTACTACT | 811 |
| RESULT 9 | | | |
| CL960332 | 1404 bp DNA linear GSS 21-SEP-2000 | | |
| LOCUS | O61FCC004254 Oryza sativa Express Library Oryza sativa (indica | | |
| DEFINITION | cultivar-group) genomic, genomic survey sequence. | | |
| ACCESSION | CL960332 | | |
| VERSION | CL960332.1 GI:52375362 | | |
| KEYWORDS | GSS. | | |
| SOURCE | Oryza sativa (indica cultivar-group) | | |
| ORGANISM | Oryza sativa (indica cultivar-group) | | |
| REFERENCE | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | |
| AUTHORS | Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza. | | |
| | 1 (bases 1 to 1404) | | |
| | Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., | | |
| | Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., | | |
| | Wong, G. K. S., Deng, X. W. and Wang, J. | | |
| | An analysis of transcriptional regulation of the rice genome and | | |
| | its comparison to Arabidopsis | | |
| | Unpublished (2004) | | |
| | Contact: Chen Chen | | |
| | Department of Bioinformatic | | |
| | Beijing Institute of Genomics | | |
| | Chinese Academy of Sciences, Beijing 101300, China | | |
| | Tel: 86-10-80481559 | | |
| | Fax: 86-10-80488676 | | |
| | Email: chenchen@genomics.org.cn | | |
| JOURNAL | | | |
| COMMENT | | | |

Db 49 GAGGATAATGATGATACGAGCAGCAGATGTCATGCCCGCTTTCGCTTCCACCCCAACT 108
 Qy 61 GluGluLeuLeuGluPheTyLeuArgArgLysValGluGlyLysArgPheAsnVal 80
 Db 109 GAAGAAGAGCTCGGGAATCTACTCTCGCGTAAGGTGGAGGGAAGCGTTTCAACGTT 168
 Qy 81 GluLeuLeuThrPheLeuAppLeuTyArgTyArgPheProTrpGluLeuProAlaMetAla 100
 Db 169 GAGCTTATTACTTCTCGATCTTATCGCTATGACCTTGGGAGCTTCCCGCTTGCGCA 228
 Qy 101 AlalLeuGlyGluLysGluTrpTyPheTyValProArgAspArgLysTyArgAsnGly 120
 Db 229 GCTATTGGTGAAGAGGTGTACTTCTATGTGCCAGAGATAGAAAGTATGCAACGGT 288
 Qy 121 AspArgProAsnArgValThrSerGlyTyTrpLysAlaThrGlyAlaAspArgMet 140
 Db 289 GATCGTCCCAATCGTGTACTACTCTGGTATTGGAAAGCAACAGAGCTGATAGATG 348
 Qy 141 IleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTySerGly 160
 Db 349 ATCCGAACCGAGAAATTTTCGCTATCGGCTCAAGAAACCCCTAGTTTCTATTCTGGG 408
 Qy 161 LysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyArgLeuProHisHis 180
 Db 409 AAAGCTCTTAAGGCATACGAACCAAGTTGGATTATGAACGAGTATCGCTTGGCCGAACAT 468
 Qy 181 GluThrGluLysTyGlnLysAlaGluLeuSerLeuCyArgValTy 196
 Db 469 GAAACTGAACGATATCAAAAGGCTGAGATATCGCTTTCGCGGTGTATC 516

RESULT 12

LOCUS CL971389 1245 bp DNA linear GSS 21-SEP-2004
 DEFINITION OaIFCC021242 Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.

ACCESSION

VERSION CL971389.1 GI:52397381

KEYWORDS

SOURCE GSS.

ORGANISM

Oryza sativa (indica cultivar-group)
 Oryza sativa (indica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

1 (bases 1 to 1245)
 Ma L., Wang C., Chen C., Liu X., Su N., Li L., Wang X., Cao M., Jiao Y., Sun N., Zhang X., Bao J., Sun D., Zhao H., Yuan L., Wong G.K.S., Deng X.W. and Wang J.

An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis

JOURNAL

Unpublished (2004)

COMMENT

Contact: Chen Chen

Department of Bioinformatic

Beijing Institute of Genomics

Chinese Academy of Sciences, Beijing 101300, China

Tel: 86-10-80481559

Fax: 86-10-80488676

Email: chenchen@genomics.org.cn

Rice genomic sequence.

Class: exon-trapped.

Location/Qualifiers

1..1245

/organism="Oryza sativa (indica cultivar-group)"

/mol_type="genomic DNA"

/db_xref="taxon:39946"

/clone_lib="Oryza sativa Express Library"

/note="Oryza sativa exon trapped genomic sequences"

ORIGIN

Alignment Scores: 1.36e-72 Length: 1245
 Pred. No.: 797.50 Matches: 185
 Score: 55.8% Conservative: 46
 Percent Similarity: 55.8% Mismatches: 88
 Best Local Similarity: 44.7%

Query Match:

DB: 39.2% Indels: 95

10 Gaps: 16

US-10-780-703-2 (1-379) x CL971389 (1-1245)

Qy 32 GlyGlnGluSerHisValGlnAsnGluAspGluAla-----AspAspHis 46
 Db 40 GGGAGGAAGAGCGCGCGCGGATAGGAGGAGCGCGCGCGCTGTGCGCGCGGACTCGCAC 99
 Qy 47 AspHisAspMetValMetProGlyPheArgPheHisProThrGluGluGluLeuGlu 66
 Db 100 GAGAACGACTGTGTATGCGGGGTTCGGTTCCACCCGAGGAGGAGGAGTATCGAG 159
 Qy 67 PheTyLeuArgArgLysValGluGlyLysArgPheAsnValGluLeuLeuLeuThrPheLeu 86
 Db 160 TTCTACTTCGCGCGGAAGGTGGAGGCGCGCGGTTCACCTCGAGCTCATCACCTCTC 219
 Qy 87 AspLeuTyArgTyArgProTrpGluLeuProAlaMetAlaAlaLeuGlyGluLysGlu 106
 Db 220 GACCTCTACCGCTTCGACCCATCGGAGCTCCCGCAATGCGGTGATTGGGGAAGAGAG 279
 Qy 107 TrpTyPheTyValProArgAspArgLysTyArgAsnGlyAspArgProAsnArgVal 126
 Db 280 TGGTTCTTTTACGTGCGCGGACCGCAAGTACCGAAGCGCCGCGCGGACCGCGGTG 339
 Qy 127 ThrThrSerGlyTyTrpLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSer 146
 Db 340 ACGGCGTCGGGTACTGGAAGCGCGGGCGCCACCGATGATCGCGCGGAGAACAGC 399
 Qy 147 ArgProIleGlyLeuLysLysThrLeuValPheTySerGlyLysAlaProLysGlyThr 166
 Db 400 CGCCCATCGGCTCAAGAAGACGCTCGTCTTCTACTCCGCGCAAGGCCCCCAAGGGCGTC 459
 Qy 167 ArgThrSerTrpIleMetAsnGluTyArgLeu-----ProHisHisGluThrGluLys 184
 Db 460 CGCAGCAGCTGGATCATGAACGAGTACCGCTTCCCCCGCGCGCGAGCGCGATCTC 519
 Qy 185 TyrGlnLysAlaGluLeuSerLeuCyArgValTyLysArgProGlyValGluAsp--- 203
 Db 520 TTCTACAGTCTGAGATCTCGTCTCGCGGTCTACAAGCGCTCCGCGCATCGACGCGC 579
 Qy 204 -----HisProSerValPro-----ArgSerLeuSerThr 213
 Db 580 CACGCGCCACCACCGCGCGCGCGCGCAACGTCGACGCGCTCGTCTCGCGCGCGCG 639
 Qy 214 ArgHisHisAsnHisAsnSerSerThrSerArgLeuAlaLeuArgGlnGlnHis 233
 Db 640 AGCGCGCGGAGGAGCACTCCGGCAACACACTCTGCTCCCGGCTTACCGGCTTCCGGCAT 699
 Qy 234 HisSerSerSerSerHisSerAspAsnAsnLeuAsnAsnAsnAsnAsnAsn 253
 Db 700 AGCGCGTACCGTCTCGTCTCGACGACGCGCGCGCGAGCACACAGTTTCCACAG 759
 Qy 254 LeuGluLysLeuSerThrGluTySerGlyAspGlySer----- 266
 Db 760 CTG-----CTCCAAGCGAGTGTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 813
 Qy 267 -----ThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 284
 Db 814 CTCCAGCATCGGCGCACGACGCGCAACAGTAAT-----GCGTCA 852
 Qy 285 GlnAsnIleTyArgProMetPro----- 292
 Db 853 CAGCTGTGTATGCGCGCGCTCCCGAAGCGCGCGCGCGCGCGCTACACGTCGCGCG 912
 Qy 293 -----TyrAspThrSerAsnAsnThrLeuIleValSerThrArgAsnHisGln----- 308
 Db 913 GCGCGCGCGCGAGAGCGCGCGTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 972
 Qy 309 -----AspAspAspGluThrAlaIleValAspAspLeuGlnArgLeuVal----- 323
 Db 973 GCGCGCGCGCGAGTAGTCTCCAGCATCGACGAGCTCAGACGCTGCTCGCGCGCGCAC 1032

Qy 324 -----AsnTyrGlnIleSerAspGlyAla----- 331
 Db 1033 AGCCATGCGCGGCTACGGTAATACCACTTGTGCGGGAAGCCATCATCTCCCTCC 1092
 Qy 332 -----ThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsnMetilePro 348
 Db 1093 TTGCGCGGCTGCACCTGATGCGGAG----- 1119
 Qy 349 AlaGlyThrIlePro-----AsnAsnAlaLeuTTPAspMet 360
 Db 1120 CTCGCGACGCTGCCCATCTCGCGCGGCTGCGCGCGCTCCGACAGGCTCTGGGAT--- 1176
 Qy 361 TTPAsnProIleValProAspGlyAsn-----ArgAspHis 372
 Db 1177 TGGAGCTCG---GTCCTCGACACGTCGACGCGCGGAGGATTAC 1215

RESULT 13
 BF625246 694 bp mRNA linear EST 18-OCT-2001
 LOCUS HVSMEO008A15f Hordeum vulgare seedling shoot EST library
 DEFINITION HVSMEO008A15f (Cold stress) Hordeum vulgare subsp. vulgare cDNA clone
 HVSMEO008A15f, mRNA sequence.
 ACCESSION BF625246
 VERSION BF625246.3 GI:1625246388
 KEYWORDS EST.
 SOURCE Hordeum vulgare subsp. vulgare
 ORGANISM Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooidae; Triticeae; Hordeum.
 REFERENCE 1 (bases 1 to 694)
 AUTHORS Wing, R., Close, T.J., Kleinohs, A., Wise, R., Begum, D., Frisch, D.,
 Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Oates, R.,
 Choi, D.W., Fenton, R.D. and Main, D.
 TITLE Development of a genetically and physically anchored EST resource
 for barley genomics: Morex cold-stressed seedling shoot cDNA
 library
 JOURNAL Unpublished (2001)
 COMMENT On Dec 18, 2000 this sequence version replaced gi:13082400.
 Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864.656.7288
 Fax: 864.656.4293
 Email: rwing@clemson.edu
 Total hq bases = 583
 Seq primer: AATTAACTCTACTAAGGG
 High quality sequence stop: 691.
 Location/Qualifiers
 1..694
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="Morex"
 /sub_species="vulgare"
 /db_xref="taxon:112509"
 /clone="HVSMEO008A15f"
 /tissue_type="Seedling shoot"
 /lab_host="TJCl21"
 /clone_lib="Hordeum vulgare seedling shoot EST library
 HVDNA0001 (Cold stress)"
 /note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;
 Seeds were surface sterilized then germinated under axenic
 conditions in the dark at room temperature on filter paper
 with water, nystatin and cefotaxime in covered
 crystallization dishes. Five-day old seedlings were
 incubated at 50C for 2 days. Shoots were then harvested,
 total RNA was prepared, poly(A) RNA was purified, one
 primary unamplified cDNA library was made, and 60000 pfu
 were in vivo excised to give pBluescript SK(-) cDNA
 phagemids. These steps were performed in the TJ Close
 laboratory at the University of California, Riverside
 (Choi, Close, Fenton). Phagemids were plated and picked at

the Clemson University Genomics Institute (CUGI) (Begum,
 Palmer, Frisch, Atkins and Wing). Plasmid DNA
 preparations, cDNA sequencing and sequence analysis were
 performed at CUGI (Wing, Yu, Frisch, Henry, Simmons,
 Oates, Rambo, Main). The sequence has been trimmed to
 remove vector sequence and contains a minimum of 100 bases
 of phred value 20 or above. For more details on library
 preparation and sequence analysis see
 http://www.genome.clemson.edu/projects/barley. To order
 this clone see http://www.genome.clemson.edu/orders Also
 see Close TJ, Wing R, Kleinohs A, Wise R (2001)
 Genetically and physically anchored EST resources for
 barley genomics. Barley Genetics Newsletter 31:29-30.
 (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"

ORIGIN

Alignment Scores:
 Pred. No.: 4,08e-72 Length: 694
 Score: 789.50 Matches: 144
 Percent Similarity: 91.8% Conservative: 12
 Best Local Similarity: 84.7% Mismatches: 11
 Query Match: 38.8% Indels: 3
 DB: 2 Gaps: 1

US-10-780-703-2 (1-379) x BF625246 (1-694)

Qy 44 AspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThrGluGluGlu 63
 Db 159 GACGCGACGAGGAGGACGCTGGTGATCCGGGTTCCCGGACCGAGGAGGAG 218
 Qy 64 LeuIleGluPheTyrLeuArgArgLysValGluGlyLysArgPheAsnValGluLeu 83
 Db 219 CTCATCAGATTCTACCTCCGCGGCAAGTGGAGGCGGAGGCTTCATGTGAGCTCATC 278
 Qy 84 ThrPheLeuAspLeuTyrArgTyrAspProTyrGluLeuProAlaMetAlaAlaGly 103
 Db 279 ACCTTCTCGACCTCTACCGCTACGACCCCTGGGAGCTCCCGAGCGTGGCAGCATTCGG 338
 Qy 104 GluLysGluTyrPheTyrValProArgAspArgLysTyrArgAsnGlyAspArgPro 123
 Db 339 GAGAAGAGGTGGTTCTTCTACGTCCGAGGACCGCAAGTACCGGACCGGGACCGGCC 398
 Qy 124 AsnArgValThrThrSerGlyTyrTyrLysAlaThrGlyAlaAspArgMetIleArgSer 143
 Db 399 AACCGGTGACGCGCTGGGGTACTGGAGGCGGCGGCGGACAGATGATCCGGGCG 458
 Qy 144 GluThrSerArgProIleGlyLeuLysThrLeuValPheTyrSerGlyLysAlaPro 163
 Db 459 GAGAGCAGCGCGCCCATCGGCTCAAGAGAGCGCTCGTCTTCTACTCCGCAAGCGGCC 518
 Qy 164 LysGlyThrArgThrSerTyrIleMetAsnGluTyrArgLeuProHisHisGluThrGlu 183
 Db 519 AAGGCGCTCCGCGACAGCTGGATCATGAACGAGTACCGGCTCCCGACCGCCGACCCGAC 578
 Qy 184 LysTyrGlnLysAlaGluIleSerLeuCysArgValTyrLysArgProGlyValGluAsp 203
 Db 579 CGATACCAACAGACAGAAATCTCTCTGCGCGCTTCTACAGCGGACCGCGGATCGACGAC 638
 Qy 204 -----HisProSerValProArgSer 210
 Db 639 GGCGCGCGCCACCCCTCTTCGCGGCGGTCTG 668

RESULT 14
 DT006901
 LOCUS DT006901
 DEFINITION VVG039B06 758717 Cabsau Cell Culture (CELu0001) Vitis vinifera cDNA
 clone VVG039B06 5, mRNA sequence.
 ACCESSION DT006901
 VERSION DT006901.1 GI:71857846
 KEYWORDS EST.
 SOURCE Vitis vinifera
 ORGANISM Vitis vinifera
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Score: 765.00 Matches: 155
Percent Similarity: 66.7% Conservative: 29
Best Local Similarity: 56.2% Mismatches: 57
Query Match: 37.6% Indels: 36
DB: 8 Gaps: 5

US-10-780-703-2 (1-379) x DR813028 (1-801)

QY 23 GlyIleGluAspAsnAspHisArgGlyGlnGluSerHisValGlnAsnGluAspGlu 42
Db 32 GCGACACCGCAGCAGCAGCAGCGCGCGCGATGGGCACCTGCGAGCAAGCGCGCGA 91
QY 43 AlaAspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThrGluGlu 62
Db 92 GGC-----GACATGGTGATGCCGGGGTTCGGTTCACCCCGCGAGGAG 136
QY 63 GluLeuIleGluPheTyrLeuArgArgValGluGlyLysArgPheAsnValGluLeu 82
Db 137 GAGCTCATCGACTTCTACCTCGCCCTAGGTGGAGGCAAGCGTTTCAACATCGAGCTC 196
QY 83 IleThrPheLeuAspLeuTyrArgTyrAspProTyrGluLeuProIleMetAlaIle 102
Db 197 ATCAACCTCGTGGACCTCTACCGCTACGACCCATGGGATCTCCCTGCTTGGCTTCGATC 256
QY 103 GlyGluLysGluTyrPheTyrValProArgAspArgLysTyrArgAsnGlyAspArg 122
Db 257 GGGGACAAGAGTGGTACTTCTACGTGGCAGAGACCGCAAGTACCGCAACGGCGACCGC 316
QY 123 ProAsnArgValThrThrSerGlyTyrTrpLysAlaThrGlyAlaAspArgMetIleArg 142
Db 317 CCCAACCGGGTCACGGCGTCCGGCTACTGGAAGGCCACGGCGCCCGCACCGCTCTAC 376
QY 143 SerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTyrSerGlyLysAla 162
Db 377 GTCGAGGTCAAGCGCGGATCGGGCTCAAGAGACGCTGGTGTCTTACGTGGGGAAGCGG 436
QY 163 ProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeuProHisGluThr 182
Db 437 CCCAAGGACTCAGGAGCAGCTGGATCATGAACGAGTACCGCTTCTTTCAGCGGAGGCC 496
QY 183 GluLysTyrGlnLysAlaGluIleSerLeuCyAsArgValTyrLysArgProGlyValGlu 202
Db 497 GACCGCTACCAAG---GAAATTTCTTTGCAAGTGTACAAACGACCGAGGATTGAA 553
QY 203 AspHisProSerValProArgSerLeuSerThrArgHisHisAsnHis--AsnSerSerTh 222
Db 554 GACAA-----CTTTCACCTCAGCACCCACCAAGATCATCTGG 594
QY 222 rSerSerArgLeuAlaLeuArgGlnGlnHisHisSerSerSerSerSerSerSerSer 239
Db 595 CTCAAAGGCTGCAGCAACCATGGAAGAGAGACACCGCAACATCGCGCTCGCCCGCAT 654
QY 240 -----HisSerAspAsnLeuAsnAsnAsnAsnAsnAsnAsnAsnAsnAsn 252
Db 655 GCGCCGCTGTTGCGAGGTGGGCTCTCATCAGTTATCATGACAAAG----- 700
QY 252 nAsnLeuGluLysLeuSerThrGluTyrSerGlyAspGlySerThrThrThrThrTh 272
Db 701 -----CCATACAGTGGAGCAACACCAACCATCTGCTGGGCATGGC 738
QY 272 rAsnSerAsnSerAspValThrIleAlaLeuAlaAsnGlnAsnIle 287
Db 739 ATCATCGCGGAGCTCGAGCAGCTACTTGGCACCACAGCGGTC 784

Search completed: April 6, 2006, 13:03:38
Job time : 4969 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocollaboration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 6, 2006, 09:49:47 ; Search time 621 Seconds
(without alignments)
4067.502 Million cell updates/sec

Title: US-10-780-703-2

Perfect score: 2034

Sequence: 1 MAIVSSTSIIPNSQVNNN.....MNPVDPGNRDHYTNIPFK 379

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abs/ABSSWEB_spool/US10780703/runat_06042006_101335_6310/app_query.fasta.1
-DB=N Geneseq -QFMT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs05p
-USER=US10780703@cgn_1_1096 @runat_06042006_101335_6310 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_21.*

1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|-------------|--------------------|
| 1 | 2034 | 100.0 | 1140 | 14 ADW23847 | Adw23847 Arabidops |
| 2 | 2034 | 100.0 | 1140 | 14 ADX17041 | Adx17041 Long Vege |
| 3 | 2006.5 | 98.6 | 1444 | 10 ADD30083 | Add30083 Plant yie |
| 4 | 2006.5 | 98.6 | 1444 | 10 ADE31520 | Ade31520 Plant yie |

| | | | | | |
|----|--------|------|------|-------------|--------------------|
| 5 | 2006.5 | 98.6 | 1444 | 12 ADI44012 | Adi44012 Plant tra |
| 6 | 1735 | 85.3 | 2606 | 14 ADW23849 | Adw23849 Arabidops |
| 7 | 1735 | 85.3 | 2606 | 14 ADX17043 | Adx17043 Long Vege |
| 8 | 781 | 38.4 | 1212 | 8 ADA70597 | Ada70597 Rice gene |
| 9 | 781 | 38.4 | 1212 | 11 ACL26812 | Acl26812 Rice ablo |
| 10 | 759 | 37.3 | 1250 | 8 ADA71117 | Ada71117 Rice gene |
| 11 | 688.5 | 33.8 | 1549 | 8 ADA70311 | Ada70311 Rice gene |
| 12 | 688 | 33.8 | 1563 | 14 ADW18167 | Adw18167 Pinus rad |
| 13 | 544.5 | 26.8 | 1158 | 8 ADA70817 | Ada70817 Rice gene |
| 14 | 536 | 26.4 | 950 | 8 ADA70350 | Ada70350 Rice gene |
| 15 | 535 | 26.3 | 1644 | 13 ADX45780 | Adx45780 Plant ful |
| 16 | 527.5 | 25.9 | 1130 | 10 ADD30816 | Add30816 Plant yie |
| 17 | 527.5 | 25.9 | 1130 | 12 ADI41856 | Adi41856 Plant tra |
| 18 | 518 | 25.5 | 1257 | 10 ADD30255 | Add30255 Plant yie |
| 19 | 518 | 25.5 | 1257 | 12 ADI44128 | Adi44128 Plant tra |
| 20 | 517.5 | 25.4 | 1152 | 8 ADA69342 | Ada69342 Rice gene |
| 21 | 517 | 25.4 | 1876 | 12 ADI42381 | Adi42381 Plant tra |
| 22 | 516 | 25.4 | 1580 | 13 ADX45919 | Adx45919 Plant ful |
| 23 | 514 | 25.3 | 830 | 8 ADA67977 | Ada67977 Arabidops |
| 24 | 514 | 25.3 | 1032 | 13 ADX30038 | Adx30038 Plant ful |
| 25 | 512 | 25.2 | 1567 | 12 ADO63463 | Ado63463 Transcrip |
| 26 | 512 | 25.2 | 1625 | 12 ADO63554 | Ado63554 Transcrip |
| 27 | 510 | 25.1 | 1236 | 12 ADO63555 | Ado63555 Transcrip |
| 28 | 509 | 25.0 | 1209 | 4 AAS00274 | Aas00274 NOR CD-11 |
| 29 | 508.5 | 25.0 | 1125 | 10 ADD30774 | Add30774 Plant yie |
| 30 | 508.5 | 25.0 | 1125 | 12 ADI41634 | Adi41634 Plant tra |
| 31 | 508.5 | 25.0 | 1125 | 14 AEA27084 | Aea27084 Stress to |
| 32 | 508.5 | 25.0 | 1326 | 3 AAC48305 | Aac48305 Arabidops |
| 33 | 508 | 25.0 | 894 | 9 ADB23141 | Adb23141 Environme |
| 34 | 508 | 25.0 | 894 | 12 ADI44144 | Adi44144 Plant tra |
| 35 | 508 | 25.0 | 894 | 12 ADO02466 | Ado02466 Thalecres |
| 36 | 508 | 25.0 | 894 | 12 ADO61748 | Ado61748 Transcrip |
| 37 | 508 | 25.0 | 894 | 14 AEA26816 | Aea26816 Stress to |
| 38 | 508 | 25.0 | 1234 | 3 AAC50693 | Aac50693 Arabidops |
| 39 | 508 | 25.0 | 1237 | 3 AAC33463 | Aac33463 Arabidops |
| 40 | 508 | 25.0 | 1260 | 10 ADD55715 | Add55715 Thalecres |
| 41 | 508 | 25.0 | 1260 | 10 ADD31002 | Add31002 Plant yie |
| 42 | 508 | 25.0 | 1260 | 12 ADI61478 | Adi61478 cdna enco |
| 43 | 507.5 | 25.0 | 903 | 3 AAC43036 | Aac43036 Arabidops |
| 44 | 507.5 | 25.0 | 1886 | 12 ADI42375 | Adi42375 Plant tra |
| 45 | 505 | 24.8 | 1551 | 13 ADX10762 | Adx10762 Plant ful |

ALIGNMENTS

RESULT 1
ADW23847
ID ADW23847 standard; cdna; 1140 BP.
AC ADW23847;
XX
AC ADW23847;
XX
DT 10-MAR-2005 (first entry)
DT
DE Arabidopsis thaliana LOV1 coding sequence, SEQ ID 1.
DE
XX
XX Flowering; plant; gene; ss.
XX
XX Arabidopsis thaliana.
OS
XX
XX Key Location/Qualifiers
XX CDS 1..1140
XX /*tag= a
XX /product= "LOV1 protein"
XX
XX KR2004075252-A.
XX
XX 27-AUG-2004.
XX
XX 20-FEB-2003; 2003KR-00010772.
XX
XX 20-FEB-2003; 2003KR-00010772.
XX
XX (UYSE-) UNIV SEOUL NAT.
PA

XX PI Ahn JH, Choi EG, Choi YD, Kim YH, Lee JS, Yoo SY;
 XX DR WPI; 2005-054880/06.
 XX DR P-PSDB; ADW23848.
 XX PT New LOV1 polynucleotide encoding a polypeptide controlling the flowering
 XX PT time of plants, i.e., either delaying or inducing early flowering of the
 XX PT plants.
 XX PS Claim 4; SEQ ID NO 1; 25pp; Korean.
 XX CC The present invention relates to the novel gene LOV1 (Long Vegetative
 CC phase 1) from Arabidopsis thaliana and its use in controlling flowering
 CC time in plants. LOV1 inhibits the flowering stimulating gene AG120. LOV1
 CC can be used for manipulating flowering time by over-expressing LOV1 to
 CC delay flowering, or inhibiting expression of LOV1 to stimulate flowering.
 CC LOV1 is also useful for identifying homologous genes in other plants. The
 CC present sequence is the LOV1 coding sequence.
 XX SQ Sequence 1140 BP; 393 A; 272 C; 216 G; 259 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3,08e-184 Length: 1140
 Score: 2034.00 Matches: 379
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 14 Gaps: 0

US-10-780-703-2 (1-379) x ADW23847 (1-1140)

QY 1 MetAlaIleValSerThrThrSerIleIleProMetSerAsnGlnValAsnAsn 20
 DB 1 ATGGCAATTGTATCTCCCAACAAGCATCATTCCTCCATGAGTAACCAAGTCAACAATAAC 60
 QY 21 GluLysGlyIleGluAspAsnAspHisArgGlyGlyGlnGluSerHisValGlnAsnGlu 40
 DB 61 GAAAGAGGTATAGAGACAAATGATCATAGAGCGGCCAAGAGAGTCATGCTCCAAATGAA 120
 QY 41 AspGluAlaAspAspHisAspMetValMetProGlyPheArgPheHisProThr 60
 DB 121 GATGAAGCTGATGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 QY 61 GluGluGluLeuIleGluPheThrLeuArgArgLysValGluGlyLysArgPheAsnVal 80
 DB 181 GAAGAAGAACTCATAGAGTTTACCTTCGCGAAAGATTGAAGGCAACCGCTTTAATGTA 240
 QY 81 GluLeuIleThrPheLeuAspLeuThrArgTyrAspProThrGluLeuProAlaMetAla 100
 DB 241 GAACATCATCACCTTCCTCGATCTTTATCGCTATGATCCTTGGGAACCTTCTCTATGGCG 300
 QY 101 AlaIleGlyGlyLysGluThrPheThrValProArgAspArgLysTyrArgAsnGly 120
 DB 301 GCGATAGAGAGAAAGAGTGGTACTTCTATGTCGCAAGAGATCGGAATATAGAAATGGA 360
 QY 121 AspArgProAsnArgValThrThrSerGlyTyrTrpLysAlaThrGlyAlaAspArgMet 140
 DB 361 GATAGACCGAACCGAGTAACGACTTCAGGATATGGAAGACCGAGCGAGCTGATAGGATG 420
 QY 141 IleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTyrSerGly 160
 DB 421 ATCAGATCGGAGACTTCTCGGCTATCGGATTAAGAAACCAACCTAGTTTCTACTCTGT 480
 QY 161 LysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeuProHisHis 180
 DB 481 AAGGCCCTTAAGGACACTGCTACTAGTATGATCATGACGATATCGTTCCTCCGACCAT 540
 QY 181 GluThrGluLysTyrGlnLysAlaGluIleSerLeuLysArgValTyrLysArgProGly 200
 DB 541 GAAACCGAGAACTACCAAAAGGCTGAATATATCATTTGCGCGAGTGTACAAAGGCCAGGA 600
 QY 201 ValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsnHisSer 220

DB 601 GTAGAGATCATCCATCGGTACGACAGCTTCTCTCTCCACAGACATCATACCATCTCA 660
 QY 221 SerThrSerArgLeuAlaLeuArgGlnGlnHisSerSerSerSerSerSerAsnHis 240
 DB 661 TCGACATCATCCGCTTTAGCCTTAAGACAAACAACACACCATTCATCTCTCTAATCAT 720
 QY 241 SerAspAsnLeuAsnAsnAsnAsnAsnAsnIleAsnAsnLeuGluLysLeuSerThrGlu 260
 DB 721 TCCGACAAACCTTTAACAACAACAACAACAACAATCAACAATCTCGAAGCTCTCCACCGAA 780
 QY 261 TyrSerGlyAspGlySerThrThrThrThrThrThrThrThrThrThrThrThrThr 280
 DB 781 TATTCCGGCGACGCGACGACAAACAACGACCAACGACGACGACGACGACGACGACGAC 840
 QY 281 AlaLeuAlaAsnGlnAsnIleTyrArgProMetProTyrAspThrSerAsnAsnThrLeu 300
 DB 841 GCTTAGCCCAATCAAAACATATATCGTCCAAATGCTTACGACACAAACGACCAACAT 900
 QY 301 IleValSerThrArgAsnHisGlnAspAspGluThrAlaIleValAspAspLeuGln 320
 DB 901 ATAGTCTCTACGAGAAATCATCAAGACGATGAGAACTGCAATGTCGATCTTCA 960
 QY 321 ArgLeuValAsnTyrGlnIleSerAspGlyAlaThrThrThrLeuMetProGlnThrGlnAla 340
 DB 961 AGACTAGTAACTACCAAAATATCAGATGGAGCGACAAAGCTAATGCCTCAAACTCAAGCG 1020
 QY 341 AlaLeuAlaMetAsnMetIleProAlaGlyThrIleProAsnAsnAlaLeuTrpAspMet 360
 DB 1021 CGCTTAGCTATGAACATGATTCCTTCGAGGAACGATTCCAACAATGCTTTGGGATATG 1080
 QY 361 TrpAsnProIleValProAspGlyAsnArgAspHisTyrThrAsnIleProPheLys 379
 DB 1081 TGGATCCATAGTACGATGAGAAACGAGATCACTATATCTATATCTTCTTTTAAAG 1137

RESULT 2
 ADX17041
 ID ADX17041 standard; cDNA; 1140 BP.
 XX AC ADX17041;
 XX DT 21-APR-2005 (first entry)
 XX DE Long Vegetative phase 1, LOV1, coding sequence, SEQ ID 1.
 XX KW Long Vegetative phase 1; LOV1; flowering; plant; gene; ss.
 XX OS Arabidopsis thaliana.
 XX FH Key Location/Qualifiers
 XX CDS 1..1140
 FT /*tag= a
 FT /product= "LOV1"
 XX US2005034194-A1.
 XX PD 10-FEB-2005.
 XX PF 19-FEB-2004; 2004US-00780703.
 XX PR 20-FEB-2003; 2003KR-00010772.
 XX PA (UYSE-) UNIV SEOUL NAT IND FOUND.
 XX PI Lee JS, Kim YH, Choi EK, Yoo SY, Ahn JH, Choi YD;
 XX DR WPI; 2005-054880/15.
 XX DR P-PSDB; ADX17042.
 XX PT New LOV1 polynucleotide encoding a polypeptide controlling the flowering
 XX PT time of plants, i.e., either delaying or inducing early flowering of the
 XX PT plants.

PT i.e. modified desirable traits, e.g. salt stress resistance or tolerance
PT to freezing.

PS Disclosure: SEO ID NO 87: 311pp: English.

xx CC The invention relates to a number of isolated cDNA sequences and their
CC encoded proteins which are especially transcription factor related cDNA/
CC and proteins. The isolated or recombinant plant transcription factor
CC polynucleotides and polypeptides are useful in producing transgenic
CC plants with commercially valuable properties, i.e. modified or altered
CC desirable traits as compared to a reference plant, e.g. salt stress
CC resistance, osmotic stress resistance, tolerance to freezing, drought,
CC low humidity tolerance, or radiation resistance. Sequence information
CC related to the polynucleotides and polypeptides can also be used in
CC bioinformatic search methods. The transgenic plant is useful for growing
CC a progeny plant from a parent plant. This sequence represents one of the
CC cDNAs of the invention

Sequence 1444 BP; 490 A; 350 C; 246 G; 358 T; 0 U; 0 Other;

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|------|
| Pred. No.: | 1,798-191 | Length: | 1444 |
| Score: | 2006.50 | Matches: | 379 |
| Percent Similarity: | 91.5% | Conservative: | 0 |
| Best Local Similarity: | 91.5% | Mismatches: | 0 |
| Query Match: | 98.6% | Indels: | 35 |
| DB: | 10 | Gaps: | 1 |

US-10-780-703-2 (1-379) x ADE31520 (1-1444)

| | | | |
|----|-----|---|-----|
| Qy | 1 | MetAlaIleValSerSerThrThrsertlleProMetSerAsnGlnValAsnAsn | 20 |
| Db | 148 | ATGGCAATTGTATCTCTCCACAAACAGCATCATTCCTCATGAGTAACCAAGTCACCAATTAAC | 207 |
| Qy | 21 | GIuLysGIyleuAspAsnAspHisAArgSGlyGIyGlnGluSerHisValGlnAsnGlu | 40 |
| Db | 208 | GAAGAAGGTATAGAGACAAATCATATAGAGCGGCCAAGAGAGTCATGTCCAAAATGAA | 267 |
| Qy | 41 | AspGluAlaAspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThr | 60 |
| Db | 268 | GATGAAGCTGATGATCATGATCATGACATGGTCATGCCCGATTTAGATTCCATCTACC | 327 |
| Qy | 61 | GIuGIuGluLeuIleGIuPheTyrlEuArgLysValGIuGIuLysAArgPheAsnVal | 80 |
| Db | 328 | GAAGAAGAACTCATAGAGTTTACCTTCGCCGAAAGTTGAAGGCCAACCGCTTTAATGTGA | 387 |
| Qy | 81 | GIuLeuIleThrPheLeuAspLeuTyrlArgTyrlAspProTrpGIuLeuProAlaMetAla | 100 |
| Db | 388 | GAACTCATCACTTTCTCGATCTTTATCGCTATGATCCTTGGGAACCTTCTGTCTATGGCG | 447 |
| Qy | 101 | AlaIleGlyGluLysGluTrpTyrlPheTyrlValProArgAspArgLysTyrlArgAsnGly | 120 |
| Db | 448 | GGCAGTAGGAGAGAAAGAGTGGTACTTCTATGTGCCAAGAGATCGGAAATATAGAAATGGA | 507 |
| Qy | 121 | AspArgProAsnArgValThrThrSerGlyTyrlTrpLysAlaThrGIuValAAspArgMet | 140 |
| Db | 508 | GATAGCCGACCGAGTAACGACTTCAGGATATTGGAAGGCCACCGAGGCTGATAGATG | 567 |
| Qy | 141 | IleArgSerGluThrSetArgProIleGlyLeuLysLysThrLeuValPheTyrlSerGly | 160 |
| Db | 568 | ATCAGATCGAGACTTCTCGGCTATCGGATTAAGAGAAACCCCTAGTTTTCTACTCTGTT | 627 |
| Qy | 161 | LysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrlArgLeuProHisHis | 180 |
| Db | 628 | AAAGCCCCCTAAAGGGCACTCGTACTAGTTGGATCATGAACGAGTATCGTCTTTCGCCACCAT | 687 |
| Qy | 181 | GIuThrGluLysTyrlGlnLysAlaGluIleSerLeuCysArgValTyrlLysArgProGly | 200 |
| Db | 688 | GAACCGAGAGTACCAAAAGGCTGGAATATCATTTGTCGGAGTGTCAAAGGCCAGGA | 747 |
| Qy | 201 | ValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsnHisAsnSer | 220 |
| Db | 748 | GTAGAAGATCATCTCCGTATCCAGTATCCAGTCTCTCTCCACAGACATCATTAACATACTCA | 807 |

PA (JIAN/) JIANG C.
 PA (HEAR/) HEARD J. E.
 PA (HAAR/) HAAKE V.
 PA (CREE/) CREELMAN R. A.
 PA (RATC/) RATCLIFFE O.
 PA (ADAM/) ADAM L. J.
 PA (REUB/) REUBER T. L.
 PA (KEDD/) KEDDIE J.
 PA (BROU/) BROUN P. E.
 PA (PILG/) PILGRIM M. L.
 PA (DUBE/) DUBELL A. N.
 PA (PINE/) PINEDA O.
 PA (YUGG/) YU G.
 XX
 PI Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;
 PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddle J, Broun PE;
 PI Pilgrim ML, Dubell AN, Pineda O, Yu G;
 XX
 DR WPI: 2004-132245/13.
 DR P-PSDB: ADI44013.
 XX
 PT New transgenic plant comprising a recombinant polynucleotide of any one
 PT of more than 500 nucleotide sequences, useful in bioinformatic search
 PT methods.
 XX
 PS Disclosure; SEQ ID NO 2475; 435pp; English.
 XX
 CC The invention describes a transgenic plant comprising a recombinant
 CC polynucleotide of any one of more than 500 nucleotide sequences fully
 CC defined in the specification or its complement. The method of the
 CC invention can be used to produce a plant having altered traits such as:
 CC enhanced tolerance to abiotic stress; glyphosate tolerance; hormone
 CC sensitivity; disease resistance; sugar sensing; early or late flowering;
 CC altered flower structure, change in stem bifurcations, altered branching
 CC pattern, reduced apical dominance, reduced trichome density; lack of
 CC trichomes; reduced ectopic trichome development; altered trichome
 CC development; increase in trichome number; altered stem morphology;
 CC increased root growth; increased root hairs; altered seed development;
 CC altered cell proliferation or cell differentiation; rapid development;
 CC premature senescence; increased necrosis; increase in seedling or plant
 CC size; decreased plant size; leaf morphology; seed morphology; seed
 CC biochemistry; increase in root ancyonins; increase in plant
 CC anthocyanins; or alteration in light response or shade avoidance. The
 CC transgenic plant, polynucleotides and polypeptides are useful in
 CC bioinformatic search methods. This sequence represents a plant
 CC transcription factor related polynucleotide.
 XX
 SQ Sequence 1444 BP; 490 A; 350 C; 246 G; 358 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.79e-181 Length: 1444
 Score: 2006.50 Matches: 379
 Percent Similarity: 91.5% Conservative: 0
 Best Local Similarity: 91.5% Mismatches: 0
 Query Match: 98.6% Indels: 35
 DB: 12 Gaps: 1

US-10-780-703-2 (1-379) x ADI44012 (1-1444)

QY 1 MetAlaIleValSerSerThrThrSerIleIleProMetSerAsnGlnValAsnAsnAen 20
 DB 148 ATGGCAATGTATCTCCCAACAGCATCATCCCATGAGTACCAAGTCAACATTAAC 207
 QY 21 GluLysGlyIleGluAspAsnAspHisArgGlyGlyGlnGluSerHisValGlnAsnGlu 40
 DB 208 GAAAGAGGTATAGAGACATCATCATGAGCGGCCCAAGAGAGTCAATGTCACAAATGAA 267
 QY 41 AspGluAlaAspAspHisAspMetValMetProGlyPheArgPheHisProThr 60
 DB 268 GATGAAGTGTATCATCATCATCATGATGATGATGATGATGATGATGATGATGATGATGAT 327
 QY 61 GluGluGluLeuIleGluPheThrLeuArgArgLysValGluGlyLysArgPheAsnVal 80

328 GAAGAGAACTCATAGAGTTTACCTTCGCCGAAAAGTTGAAGGCAAAACGCTTTAATGTA 387
 QY 81 GluLeuIleThrPheLeuAspLeuTyrArgTyrAspProThrGluLeuProAlaMetAla 100
 DB 388 GAACTCATCTTTCTCCGATCTTATCGGTATATCTCTTGGGAACATTCCTGCTATGGC 447
 QY 101 AlaIleGlyGluLysGluTyrPheTyrValProArgAspArgLysTyrArgAsnGly 120
 DB 448 GCGATAGAGAGAAAGAGTGTACTTCTATGTGCGCAAGAGATCGGAATATAGAAATGGA 507
 QY 121 AspArgProAsnArgValThrThrSerGlyTyrTrpLysAlaThrGlyAlaAspArgMet 140
 DB 508 GATAGACCGAACCGAGTAACGACTTCAGGATATTTGGAAAGCCACCGGAGCTGATAGGATG 567
 QY 141 IleArgSerGluThrSerArgPheIleGlyLeuLysValThrLeuValPheTyrSerGly 160
 DB 568 ATCAGATCGGAGACTTCTCGGCTATCGGATTAAGAAACCCCTAGTTTCTACTCTGCT 627
 QY 161 LysAlaProLysGlyThrArgThrSerTrpIleMetAsnGluTyrArgLeuProHisHis 180
 DB 628 AAGCCCTTAAGGCACTCGTACTAGTTGGATCATGAAGAGTATCGTCTTCGCAACCAT 687
 QY 181 GluThrGluLysTyrGlnLysAlaGluIleSerLeuCysArgValTyrLysArgProGly 200
 DB 688 GAAACCGAGAAGTACCAAAAGGCTGAAATATCATTTGTGCCGAGTGTACAAAAGGCCAGGA 747
 QY 201 ValGluAspHisProSerValProArgSerLeuSerThrArgHisHisAsnHisAsnSer 220
 DB 748 GTAGAAATCATCTCCGTCACCACTGTTCTCTCCACAGACATCATTAACATTAACATCA 807
 QY 221 SerThrSerArgLeuAlaLeuArgGlnGlnGlnHisHisSerSerSerSerSerAsnHis 240
 DB 808 TCGACATCATCTCCGTTAGCCTTAAGACACACACACACATTCATCTCTCTCTAATCAT 867
 QY 241 SerAspAsnLeuAsnAsnAsnAsnAsnLeuLeuLeuLeuLeuLysLeuSerThrGlu 260
 DB 868 TCCGACAAACACCTTAACACACACACACATCAACATCTCGAGAAGCTCTCCACCCGAA 927
 QY 261 TyrSerGlyAspGlySerThrThrThrThrThrThrThrThrThrThrThrThrThrThr 280
 DB 928 TATTCGCGGACGCGGACAT 987
 QY 281 AlaLeuAlaAsnGlnAsnIleTyrArgProMetProTyrAspThrThrThrThrThrThr 300
 DB 988 GCTCTAGCAATCAAAACATATATCGTCCATGCTTACGACACACACACACACACACAT 1047
 QY 301 IleValSerThrArgAsnHisGlnAspAspGluThrAlaIleValAspLeuGln 320
 DB 1048 ATAGTCTCTACGAGAAATCATCAAGACGATGATGAACTGCCATTTGTTGACGATCTTCAA 1107
 QY 321 ArgLeuValAsnTyrGlnIleSerAspGly----- 330
 DB 1108 AGACTAGTTAACTACCAAAATATCATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1167
 QY 330 ----- 330
 DB 1168 GCTCAACAGTTTTCATCATCTACTCAACACAAATGCTACGCAAAAGCGCATTAACAATTGGTG 1227
 QY 331 -----AlaThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsn 345
 DB 1228 GCTGCGGCGACTACAGCGCACACGCTAATGCTCAAACTCAAGCGCGCTTAGCTATGAAC 1287
 QY 346 MetIleProAlaGlyThrIleProAsnAsnAlaLeuTyrAspMetTyrAsnProIleVal 365
 DB 1288 ATGATTCCTCGAGAACGATTCCAAAACAAATGCTTTGTGGGATATGTGAATCCAATAGTA 1347
 QY 366 ProAspGlyAsnArgAspHisTyrThrAsnIleProPheLys 379
 DB 1348 CCAGATGGAACACAGATCATCTATCATTAATATTCCTTTTAAG 1389

RESULT 6
 ADW23849
 ID ADW23849 standard; DNA; 2606 BP.

XX AC ADW23849;
 XX DT 10-MAR-2005 (first entry)
 XX DE Arabidopsis thaliana LOV1 genomic sequence, SEQ ID 3.
 XX KW Flowering; plant; gene; ds.
 XX OS Arabidopsis thaliana.
 XX FH Key Location/Qualifiers
 XX FT CDS 1..2606
 XX FT /*tag= b
 XX FT /product= "LOV1 protein"
 XX FT exon 1..292
 XX FT /*tag= a
 XX FT /number= 1
 XX FT intron 293..622
 XX FT /*tag= c
 XX FT /number= 1
 XX FT exon 623..892
 XX FT /*tag= d
 XX FT /number= 2
 XX FT intron 893..1923
 XX FT /*tag= e
 XX FT /number= 2
 XX FT /cons_splice= (5'site:NO,3'site:NO)
 XX FT exon 1924..2351
 XX FT /*tag= f
 XX FT /number= 3
 XX FT intron 2352..2456
 XX FT /*tag= g
 XX FT /number= 3
 XX FT /cons_splice= (5'site:NO,3'site:NO)
 XX FT exon 2457..2606
 XX FT /*tag= h
 XX FT /number= 4
 XX KR2004075252-A.
 XX 27-AUG-2004.
 XX 20-FEB-2003; 2003KR-00010772.
 XX 20-FEB-2003; 2003KR-00010772.
 XX (UYSE-) UNIV SEOUL NAT.
 XX Ahn JH, Choi EG, Choi YD, Kim YH, Lee JS, Yoo SY;
 XX WPI; 2005-054880/06.
 XX P-PSDB; ADW23848.
 XX New LOV1 polynucleotide encoding a polypeptide controlling the flowering
 XX time of plants, i.e., either delaying or inducing early flowering of the
 XX plants.
 XX Claim 5; SEQ ID NO 3; 25pp; Korean.
 XX The present invention relates to the novel gene LOV1 (Long Vegetative
 XX phase 1) from Arabidopsis thaliana and its use in controlling flowering
 XX time in plants. LOV1 inhibits the flowering stimulating gene AG120. LOV1
 XX can be used for manipulating flowering time by over-expressing LOV1 to
 XX delay flowering, or inhibiting expression of LOV1 to stimulate flowering.
 XX LOV1 is also useful for identifying homologous genes in other plants. The
 XX present sequence is the LOV1 genomic sequence.
 XX Sequence 2606 BP; 896 A; 509 C; 388 G; 813 T; 0 U; 0 Other;
 XX
 XX Alignment Scores:
 XX Pred. No.: 3.28e-155 Length: 2606
 XX Score: 1735.00 Matches: 378

Percent Similarity: 43.5% Conservativity: 0
 Best Local Similarity: 43.5% Mismatches: 1
 Query Match: 85.3% Indels: 490
 DB: 14 Gaps: 3
 US-10-780-703-2 (1-379) x ADW23849 (1-2606)
 QY 1 MetAlaIleValSerSerThrThrSerIleProMetSerAenGlnValAsnAsn 20
 DB 1 ATGGCAATTGTATCTCCACAAAGCATCATCTCCATGAGTAACCAAGTCAACAATAAC 60
 QY 21 GluLysGlyIleGluAspAsnAspHisArgGlyGlnGlnGluSerHisValGlnAsnGlu 40
 DB 61 GAAAAGGTATAGAGACAATGATCATAGAGCGGCCCAAGAGAGTCATGTCCAAAATGAA 120
 QY 41 AspGluAlaAspAspHisAspHisAspMetValMetProGlyPheArgPheHisProThr 60
 DB 121 GATGAAGCTGATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 QY 61 GluGluGluLeuIleGluPheTyrLeuArgArgGlyValGluGlyLysArgPheAsnVal 80
 DB 181 GAAGAAGACTCATAGAGATTTTACCTTCGCCGAAAAGTTGAAGCAAAACGGCTTTAATGTA 240
 QY 81 GluLeuIleThrPheLeuAspLeuTyrArgTyrAspProTyrGluLeuPro----- 97
 DB 241 GAATCATCACTTCTCTCGATCTTATCGTATGATCCTTGGGAACCTTCTGTAATATAT 300
 QY 97 ----- 97
 DB 301 ACATTCACATAAAACACACATAAATCATCTCAAACTATTGGAAATCTTAATTTCTATTCA 360
 QY 97 ----- 97
 DB 361 TATGTTAAGATCTTCT 420
 QY 97 ----- 97
 DB 421 ATATATGTACCTACCT 480
 QY 97 ----- 97
 DB 481 CTTTAAACGCTTCTCCCTCTTCT 540
 QY 97 ----- 97
 DB 541 TATATATCATAGATATACAAATATGTGTATGTATGTATGTATGTATGTATGTATGTATGTAT 600
 QY 98 -----AlaMetAlaAlaIleGlyGluLysGluTyrPheTyr 110
 DB 601 TCATAGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 QY 111 ValProArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGly 130
 DB 661 GTGCCAAGAGATCGGAAATATAGAAATGGAGATAGACCGAACCGAGTAACGACTTCAGGA 720
 QY 131 TTTTTPLeAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGly 150
 DB 721 TATGGAAAGCCACCGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 QY 151 LeuLysLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTyr 170
 DB 781 TTAAAGAAACCTTAGTTTCT 840
 QY 171 IleMetAsnGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLysAla----- 188
 DB 841 ATCATGAACGAGTATCGTCTTCCGCCACCATGAACCGAGAAGTACCAAAAGGT-ATAAAT 899
 QY 188 ----- 188
 DB 900 TCTACTATAACTCTATATATATATCTCTATTCATACATAGATATAACCCCTAGTAGGTG 959
 QY 188 ----- 188

Db 121 GACCCCTGGAGCTCCCGGAATGAGCATCTAATAACAAGCTCGATGGGGTGCAGCA 180
Qy 99 MetAlaAlaIleGlyGluLysGluTyrPheTyrValProArgAspArgLysTyrArg 118
Db 181 ATGGCGGGATAGGGGAGAGAGTGGTCTTCTAGTGCCTCGGAGCAGAGATACAGG 240
Qy 119 AenGlyAspArgProAsnArgValThrThrSerGlyTyrTTPLeAlaThrGlyAlaAsp 138
Db 241 AACGGGACCGGCGCAACCGGGTACGGGTCTGGAAGCGGACGGGGCGCGAC 300
Qy 139 ArgMetIleArgSerGluThrSerArgProIleGlyLeuLysLysThrLeuValPheTyr 158
Db 301 CGGATGATCCGAGCGAGAACACCGCCCATCGGGCTCAAGAACACGCTTGTCTCTAC 360
Qy 159 SerGlyLysAlaProLysGlyThrArgThrSerTTPLeMetAsnGluTyrArgLeuPro 178
Db 361 TCCGCAAGGCCCGCCAGGGGTCCGACGAGCTGATCATGACAGATATCGCCTTCCC 420
Qy 179 HisHisGluThrGluLysTyrGlnLysAlaGluLeuSerLeuCybArgValTyrLysArg 198
Db 421 CCGCGCGACACCGCGCTACCAAGACTGAAATCTCACTATGCGCGTGTACAAAGCGC 480
Qy 199 ProGlyValGluAsp---HisProSerVal-ProArg-----Se 210
Db 481 ACCGCGATCGACGACGGCGCTATCCAGCGCGGTCTCCGCGCGCACTCCCGC 540
Qy 210 rLeuSerThrArgHisHisAsnHisAsnSerSerThrSerArgLeuAlaLeuArgG1 230
Db 541 GCGCGCGCGCGCGACCGGTACAGAACACAGCAAGGCTCTCTGTCAGCTCCACGCGC 600
Qy 230 n-----GlnGlnHisHisSerSerSerSerSerSerSerSerSerSerSerSer 247
Db 601 ACGCGCGCTCCACTCCGCTCCAGCTCCACTCTCCAGCGAGTGCAGCTCGCGCGCC 660
Qy 247 nAsnAsnAsnLeuAsnAsnLeuGluLysSerThrGluTyrSerGlyAspGlySerTh 267
Db 661 GCCATCGTCGCGGACCGCGCCCATGGTGGCGCACAGGCGACCGCTCCCGCGCGC 715
Qy 267 rThrThrThrThrThrAsnSerAsnSerAspValThrIleAlaLeuAlaAsnGlnAsnI1 287
Db 716 -ACCATCAGCAGCAGCAGCTCCACGCTCCAGCTTCGCGGGTACTCGCAGAAAT 774
Qy 287 eTyrArg-----ProMetProTyrAspThrSerSerAsnAsnThrLeuIleValSerThrAr 305
Db 775 TCCTCGATGGCTTCGCTCGCGCGCGCGATCAGCAGCAGCAATTTCAACAGACTTCGCG 834
Qy 305 g----- 305
Db 835 GCGGCATTGTACAGCAGTACTCCAAAGAACACAGCGGTGCTTCGCTCCACGTAATCG 894
Qy 306 -----AsnHisGlnAspAspAspGluThrAlaIleValAspAspLe 319
Db 895 CTGCTCAACCTCGTCAACGCGGCATCCATG-GGCAGCTCCGCGCTGCATGTGATGAGCT 953
Qy 319 uGlnArgLeuValAsnTyr-----GlnIleSerAspGlyAlaThrThr----- 333
Db 954 GAGCTCGTGTGGCCAGCGCACCGCTCTACATTAACCCCGCGCGCGCGCACCA 1013
Qy 334 -----LeuMetProGlnThrGlnAlaAlaLeuAlaMetAsnMetIlePr 348
Db 1014 CTATAGTCAATTCCTTCACTCTACGACGCGGTGCTGCATCAACCAACGCGCGCGC 1073
Qy 348 aAlaGly-----ThrIlePro-----As 354
Db 1074 GCTCGGACGACGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1133
Qy 354 nAsnAlaLeuTyrAspMetTyrAsnProIleValProAspGlyAsnArgAspHisTyrTh 374
Db 1134 CGACAGATCTGGGACTGGATTAATCCGATCCCGAGCGCGAGTAGGATTTACAGCAC 1193
Qy 374 rAsnIleProPheLys 379
Db 1194 CAGTACTGATTCAG 1209

RESULT 9

ACL26812
ID ACL26812 standard; cDNA; 1212 BP.

XX ACL26812;

AC 02-JUN-2005 (first entry)

XX Rice abiotic stress responsive polynucleotide SEQ ID NO:768.

XX ss; abiotic stress tolerance; transgenic plant; plant; cereal;
KW agriculture.

XX Oryza sativa.

XX WO2003008540-A2.

XX 30-JAN-2003.

XX 21-JUN-2002; 2002WO-US019668.

XX 22-JUN-2001; 2001US-0300112P.

XX 24-AUG-2001; 2001US-0314662P.

XX 26-SEP-2001; 2001US-0325277P.

XX 21-NOV-2001; 2001US-0332132P.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;
PI Moughamer T, Provart N, Rickes D, Zhu T;
XX WPI; 2003-248011/24.

XX New stress-responsive nucleic acid, useful for altering the
responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
stress, salt stress or osmotic stress.

XX Claim 1; SEQ ID NO 768; 89pp; English.

XX The invention relates to novel abiotic stress responsive polynucleotides
and polypeptides. Also disclosed are vectors, expression cassettes, host
cells, and plants containing such polynucleotides. Also disclosed are
methods for using the polynucleotides and polypeptides to alter the
responsiveness of a plant to abiotic stress. The invention is useful in
agriculture. The nucleic acid is useful for determining whether a test
plant has been exposed to an abiotic stress condition. It is also useful
for selecting an agent that alters abiotic stress regulated
polynucleotide expression in a plant cell, and to identify a homolog or
ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
molecule and the polypeptide encoded by it are useful in altering the
responsiveness of a plant to an abiotic stress, such as cold stress, salt
stress, osmotic stress or any of their combinations. The present sequence
is used in the exemplification of the invention

XX Sequence 1212 BP; 245 A; 421 C; 361 G; 185 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.33e-64 Length: 1212
Score: 781.00 Matches: 185
Percent Similarity: 57.1% Conservative: 47
Best Local Similarity: 45.6% Mismatches: 94
Query Match: 38.4% Indels: 81
DB: Gaps: 11

US-10-780-703-2 (1-379) x ACL26812 (1-1212)

Qy 52 MetProGlyPheArgPheHisProThrGluGluLeuIleGluPheTyrLeuArgArg 71

Db 1 ATGCCCCGGTTTCAGTTTCCACCCCGAGGAGGAGTGTATCGATTCTACCTCGCGCGG 60

Qy 72 LysValGluGlyLysArgPheAsnValGluLeuIleThrPheLeuAspLeuTyrArgTyr 91

Db 61 AAGGTGGAGGCAAGCGCTTCAACGTCGAGCTCATCACCTTCTCGACCTCTACCGCTAC 120
Qy AspProTTPGlulLeuPro-----Ala 98
Db 121 GACCCCTGGAGCTCCCGGAATTGAGCATCTAATTAAAGCTCGATGGGGTGCAGCA 180
Qy MetAlaAlaileGlyLulysGluTTPtyrPheTyrValProArgAspArgLysTyrArg 118
Db 181 ATGGCGCGCATAGGAGAGAGAGTGTCTTCTAGTCCCTCGGACAGGAGTACAGG 240
Qy AsnGlyAspArgProAsnArgValThrThrSerGlyTyrTriPlysAlaThrGlyAlaAsp 138
Db 241 AACGGGACCGGCGGAACCGGTCGCGCTCGGGTACTGGAAGCGACGCGGGCGGAC 300
Qy ArgMetIleArgSerGluThrSerArgProIleGlyLeuLysIleThrLeuValPheTyr 158
Db 301 CGGATGATCGAGCGGAGACCAACCGCCCATCGGGCTCAAGAAGACGCTTGTCTTCTAC 360
Qy SerGlyLysAlaProLysGlyThrArgThrSerTriPileMetAsnGluTyrArgLeuPro 178
Db 361 TCGGCAAGCCCGCCCAAGGGCTCCGACGAGCTGGATCATGACGATTCGCTTCC 420
Qy HisHisGluThrGluLysTyrGlnLysAlaGluIleSerLeuCysArgValTyrLysArg 198
Db 421 CCGCGCACCGACCGCTACCAACAGACTGAAATCTCACTATGCCGCGGTGTACAAGCGC 480
Qy ProGlyValGluAsp---HisProSerVal-ProArg-----Se 210
Db 481 ACCGGCATCGACGACCGCATGCCAGGTATCCACGGCGCGGTGTCGGCGCACTCCCGC 540
Qy rLeuSerThrArgHisHisAsnHisAsnSerSerThrSerArgLeuAlaLeuArgG1 230
Db 541 GCGGCGCGCGCGCACCGGTACAGGACAAACGAAGCTCTCTGTCGAGCTCCAGCGCG 600
Qy n-----GlnGlnHisHisSerSerSerSerSerSerSerSerSerSerSerSerSer 247
Db 601 ACGCGCGCTCCAACTCCGTCGAAGTCCACCTCTCAGCAGCGAGTGCAGCTCGCGCGC 660
Qy nAsnAsnAsnIleAsnAsnLeuGluLysLeuSerThrGluTyrSerGlyAspGlySerTh 267
Db 661 GCCATCGTCACGACACCGCCCATGTGTGGCGCACAAAGCGCACCGTCGCGCGCGC 715
Qy rThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 287
Db 716 -ACCATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 774
Qy eTyrArg-----ProMetProTyrAspThrSerAsnAsnThrLeuIleValSerThrAr 305
Db 775 TCTCGATGGCTTCGCTCGCGCGCGCGGATCAGCAGCAATTTTCAGCAAGACTTCGCG 834
Qy g----- 305
Db 835 GCGGCACTGTACGACGAGTACTCAAGAACACGAGCGGTGCTTCCGCTCCAGTACTCG 894
Qy 306 -----AsnHisGlnAspAspAspGluThrAlaIleValAspAspIle 319
Db 895 CTGCTCAACCTCGTCAACGCGGATCCATG-GGCAGCTCCGCGCTGCCATGATGAGCT 953
Qy uGlnArgLeuValAsnTyr-----GlnIleSerAspGlyAlaThrThr----- 333
Db 954 GAGCTCGTGTGGCGCACCGGACGCGCTCTACATTAAACCGCGCGCGCGGCGGACCA 1013
Qy 334 -----LeuMetProGlnThrGlnAlaAlaLeuAlaMetAsnMetIlePr 348
Db 1014 CTATAGTCAATTCCTTTCACCTGCTAGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1073
Qy oAlaGly-----ThrIlePro-----As 354
Db 1074 GCTCGGGACGACGACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1133
Qy nAsnAlaLeuTTPAspMetTrpAsnProIleValProAspGlyAsnArgAspHisTyrTh 374
Db 1134 CGACAGGATCTGGGATCTGGNATTAATTCGATCCCGGAGCGCGGAGGTAGGGATTACAGC 1193

Qy 374 rAsnIleProPheLys 379
Db 1194 CAGTACTGGATTCAAG 1209

RESULT 10

ADA71117

ID ADA71117 standard; DNA; 1250 BP.

XX

AC ADA71117;

XX

DT 20-NOV-2003 (first entry)

XX

DE Rice gene, SEQ ID 4440.

XX

KW Plant; bacterial infection; fungal infection; viral infection; rice;

KW gene; ds.

XX

OS Oryza sativa.

XX

PN WO2003000898-A1.

XX

PD 03-JAN-2003.

XX

PF 22-JUN-2001; 2001WO-IB001105.

XX

PR 22-JUN-2001; 2001WO-IB001105.

XX

PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX

PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX

DR WPI; 2003-175290/17.

XX

PT Identifying at least one gene involved in plant resistance or response to

PT pathogenic infection for conferring resistance or tolerance to a plant to

PT bacterial, fungal or viral infection by determining or detecting plant

PT gene expression.

XX

PS Claim 6; SEQ ID NO 4440; 899pp; English.

XX

CC The present invention relates to a method (M1) for identifying genes

CC involved in plant resistance or response to pathogenic infection. M1

CC comprises identifying a gene whose expression is significantly altered in

CC the incompatible interaction of plant gene expression relative to

CC expression of the gene in an uninfected plant, in a mutant plant that

CC does not express a gene associated with response to pathogenic infection,

CC or in a corresponding incompatible or compatible interaction. (M1) is

CC useful for conferring resistance to resistance or tolerance to a plant to

CC bacterial, fungal or viral infection. The present sequence was used to

CC illustrate the invention.

XX

SQ Sequence 1250 BP; 291 A; 388 C; 339 G; 230 T; 0 U; 2 Other;

XX

Alignment Scores:

Pred. No.: 1,748-62 Length: 1250

Score: 759.00 Matches: 174

Percent Similarity: 50.8% Conservative: 45

Best Local Similarity: 40.4% Mismatches: 72

Query Match: 37.3% Indels: 140

DB: 8 Gaps: 14

US-10-780-703-2 (1-379) x ADA71117 (1-1250)

Qy 50 MetValMetProGlyPheArgPheHisProThrGluGluLeuLeuPheTyrLeu 69

Db 1 ATGGTAATGCCGGGTTCCGGTTCCACCCGAGGAGGAGCTCATCGAGTTCTACCTC 60

Qy 70 ArgArgLysValGluGlyLysArgPheAsnValGluLeuIleThrPheLeuAspLeuTyr 89

Db 61 CGCGGAAAGGTGGAGGGGAAACGGTTCAACATCGAGCTCATCGCTTCTCGACCTCTAC 120


```
Qy 90 ArgTyrAspProTyrGluLeuProAlaMetAlaIleGlyGluLysGluTyrPhe 109
Db 121 CGCTATGACCGTGGATCTTCCGCTCTGGCTCAATTGGGACAAGAGTGGTCTTC 180
Qy 110 TyrValProArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrSer 129
Db 181 TATGTGCCAAGGACCGCAAGTACGAAACGCGGATCGGCCCAACCGAGTACGCGCATCG 240
Qy 130 GlyTyrTrpLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIle 149
Db 241 GGGTACTGGAGGCCACAGGGCGGATAGTAGTGGTAGAGGAGTACCGTCTCATC 300
Qy 150 GlyLeuLysLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSer 169
Db 301 GGCCTCAAGAGACGCTCGTCTTCTACGTGCGCAAGGCGCCCAAGGGGCTTCGCAGCAGC 360
Qy 170 TrpIleMetAsnGluTyrArgLeuProHisHisGluThrGluLysTyrGlnLys 187
Db 361 TGGATCATGAACGAGTATCGTCTCCCTCATGGCGACGCTGACCGCTACCAAAAGTAGCT 420
Qy 187 ----- 187
Db 421 TTTGATATAGTACCGGCGAGGAGCAAGCATATGCTACTGCTATGCTATTGGCAGCTCT 480
Qy 188 -----AlaGlu 189
Db 481 TGTCTCTCAGCTGGTGATGAATTAAGTACACACCATCGGAATTTGTATGTAGAGAA 540
Qy 190 IleSerLeuCysArgValTyrLysArgProGlyValGluAspHis 204
Db 541 ATTTCTGTGTGTGAGTCTACAAACGCCAGGATCGAAGCAACTCCACCTVACCGGC 600
Qy 205 -----ProSerValProArgSerLeuSerThrArgHis 215
Db 601 ACAACCACTAAATCTCCGGCTCAAGGCGCGCGATGGGCAAGCAAGCAGCGCGMC 660
Qy 216 HisAsnHisAsnSerThrSerArgLeuAlaLeuArgGlnGlnGln 232
Db 661 CGAACCGGAGTCTATCAGCGGCGGCTGGGCGCCCGCTCGCCCGATGTTGAC 720
Qy 233 -----HisHisSerSer 241
Db 721 GCGCGCGCGCGGCGCACCTCATCTCCGCTCATGGGAAGGCAACATACACGCTCA 780
Qy 242 AsnAsnAsnLeuAsnAsnAsnAsn 251
Db 781 CACACCGCGCGGAGGAGCAAGCCCATCGCCATGTCCTGCGCGCGCGCGCGCTG 840
Qy 252 -----AsnAsnLeuGluLys 256
Db 841 CATCTCTCGGCGCGCGGATGTTCCGTCGACGGCGTGGTGGCTCGCTGAGCTCCAC 900
Qy 257 LeuSerThrGluTyrSerGlyAsp 269
Db 901 ACGTCGAGGAGGAGGAGCGACGCGCTCTTCCACCTCAAGCGCGCGCGCGACGACG 960
Qy 270 ThrThrThrAsnSerAsnSerAspValThrIleAlaLeuAlaAsnGlnAsnIleTyrArg 289
Db 961 AACACGAGC -----ACGATGTCATCATCCATCATCATCATCAG 996
Qy 290 ProMetProTyrAspThr 304
Db 997 CAGTGTCTGCTTCACTCATGCTCTGCTCAACACCAACTGTCAGCAATGGCG 1050
Qy 305 ArgAsnHisGlnAspAspGluThrAlaIleValAspAspLeuGlnArgLeuVal 323
Db 1051 -----ACCATCCCATCGAGAGTGAAGTGGCGGATTTGA 1086
Qy 324 -----AsnTyrGlnIleSerAspGlyAlaThrThrLeu 334
Db 1087 TCTTACAGCAACATCCACCATCAACCAAACTCTACTACAACTCATCAATCAACTGCTTTG 1146
Qy 335 MetProGlnThrGlnAlaAlaLeuAlaMetAsnMetIleProAlaGlyThrIleProAsn 354
```

```
Db 1147 CCGCCATCGCAACAAGCCCTCTT-----CTTCCATTGCTAGCATG----- 1188
Qy 355 AsnAlaLeuTrpAspMetTrpAsnProIleVal 365
Db 1189 GAGAAGATTGGGAT---TGGAACCTCTCTCTA 1218
RESULT 11
ADA70311
ID ADA70311 standard; DNA; 549 BP.
XX
AC ADA70311;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 3634.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX
OS Oryza sativa.
XX
PN WO2003000898-A1.
PD
PP 03-JAN-2003.
PP
PP 22-JUN-2001; 2001WO-IB001105.
PP
PR 22-JUN-2001; 2001WO-IB001105.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
DR WPI; 2003-175290/17.
PT
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
PS Claim 6; SEQ ID NO 3634; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
SQ Sequence 549 BP; 103 A; 174 C; 194 G; 78 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3.05e-56 Length: 549
Score: 688.50 Matches: 125
Percent Similarity: 88.8% Conservative: 10
Best Local Similarity: 82.2% Mismatches: 12
Query Match: 33.8% Indels: 5
DB: 8 Gaps: 1
US-10-780-703-2 (1-379) x ADA70311 (1-549)
Qy 32 GlyGlnGluSerHisValGlnAsnGluAspGluAla-----AspAspHis 46
Db 49 GGGAGAGAAAGCGGCGCGCGGATCAGGAGAGGCGCGCGCTGTCGCGGCGGACTCGCAC 108
Qy 47 AspHisAspMetValMetProGlyPheArgPheHisProThrGluGluLeuIleGlu 66
-----
```

109 GAGAACGACCTGTGTATGTCGGGGTTCGGTCCACCCGAGGAGGAGGAGCTGATCGAG 168
 67 PheTyrLeuArgArgLysValGluGlyLysArgPheAsnValGluLeuLeuThrPheLeu 86
 169 TTCCTACCTCCGGGAGGTGGAGGCGCGGTTCAACGTGAGCTCATCCTTCCTC 228
 87 AspLeuTyrArgTyrAspProTTPGluLeuProAlaMetAlaAlaGlyGluGlyGlu 106
 229 GACCTTACCGTTCACCCATCGGAGCTCCCGGCATGCGGTGATAGGGGAGAGAG 288
 107 TrpTyrPheTyrValProArgAspArgLysTyrArgAsnGlyAspArgProhenArgVal 126
 289 TGGTCTTCTACGTGCGGGACCGCAAGTACCGGACCGGACCGGCGGACCGGGTG 348
 127 ThrThrSerGlyTyrTrpLysAlaThrGlyAlaAspArgMetIleArgSerGluThrSer 146
 349 ACGGCGTGGGGTACTGGAGGCGCGGGCGCGCGGATGATCCGCGCGGAGAACAGC 408
 147 ArgProIleGlyLeuLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThr 166
 409 CGCCCATCGGCTCAAGAGAGCGCTCTTCTACTCCGGCAGGCCCCCAGGGGCTC 468
 167 ArgThrSerTrpIleMetAsnGluTyrArgLeuPro 178
 469 CGCAGCAGCTGGATCATGAACGAGTACCGCTTCCCC 504

RESULT 12

ADW18167

ID ADW18167 standard; cDNA; 1563 BP.

XX AC ADW18167;

XX DT 24-MAR-2005 (first entry)

DE Pinus radiata transcription factor cDNA NAC family Seq 1946.

KW gene; ss; plant; transcription; gene regulation; gene expression;
 KW transgenic plant; drought resistance; disease resistance; salt tolerance;
 KW cold tolerance; freezing tolerance; flowering; flavor enhancer;
 KW flower color.

XX OS Pinus radiata.

XX PN WO2005001050-A2.

XX PD 06-JAN-2005.

XX PF 07-JUN-2004; 2004WO-US017965.

XX PR 06-JUN-2003; 2003US-0476189P.

XX PA (ARBO-) ARBORGEN LLC.

XX PI Blokeberg LN, Bryant C, Connert MB, Emerson SV, Frost MJ;
 PI Forster RLS, Grigor M, Higgins C, Lasham A, Lund ST, Magusin A;
 PI Phillips J, Puthgae S, Veerakone S, Westwood C, Gause K, Wood M;
 XX WPI; 2005-075542/08.
 DR P-PSDB; ADW18488.

XX PT New polynucleotides isolated from plants encoding transcription factors,
 PT and polypeptides encoded by such polynucleotides, useful for regulating
 PT gene transcription and gene expression.

XX PS Claim 3; SEQ ID NO 1946; 1265pp; English.

XX CC This invention relates to novel isolated plant nucleic acid molecules, or
 CC variants thereof, that encode transcription factors. Specifically, it
 CC refers to transcription factor proteins that are capable of binding to
 CC DNA in order to regulate gene transcription and gene expression in a
 CC plants, in particular Eucalyptus grandis and Pinus radiata. The present
 CC invention describes DNA constructs containing DNA encoding a
 CC transcription factor that regulates the promoter, which is operably

CC linked to the desired nucleic acid to be expressed. It further provides
 CC transgenic plants expressing a transcription factor that confers a trait
 CC to the plant such as increased drought, salt or disease tolerance, height
 CC change, enhanced cold/ frost tolerance, enhanced color, health and
 CC nutritional characteristics, as well as improved taste, starch
 CC composition, flower longevity and germination, amongst others.
 CC Accordingly, such plants that are successfully transformed with a DNA
 CC construct can be characterized by a difference in flower color, petal or
 CC leaf shape and size, aroma or plant height. This polynucleotide is a
 CC plant transcription factor cDNA sequence of the invention.

XX SQ Sequence 1563 BP; 501 A; 280 C; 369 G; 413 T; 0 U; 0 Other;

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 1,41e-55 | Length: | 1563 |
| Score: | 688.00 | Matches: | 160 |
| Percent Similarity: | 53.6% | Conservative: | 42 |
| Best Local Similarity: | 42.4% | Mismatches: | 119 |
| Query Match: | 33.8% | Indels: | 56 |
| DB: | 14 | Gaps: | 9 |

US-10-780-703-2 (1-379) x ADW18167 (1-1563)

| | | | |
|----|-----|---|-----|
| QY | 38 | GlnAsnGluAspGluAlaAspAsp-----HisAspHisAspMetValMetPro | 53 |
| DB | 145 | CGCAAGAGAGTCAAGGCGAGATGAAGAGCAGCTTCAAGCAACAATCAGGATGTGTCATCGC | 204 |
| QY | 54 | GlyPheArgPheHisProThrGluGluLeuLeuGluPheTyrLeuArgArgLysVal | 73 |
| DB | 205 | GGTTTCAGATTCCATCCCTACGGAAGAAGAAATGTTGGAGTTTTATCTCGCGCAAGAAGTG | 264 |
| QY | 74 | GluGlyLysArgPheAsnValGluLeuIleThrPheLeuAspLeuTyrArgTyrAspPro | 93 |
| DB | 265 | GAGGGGAAGCATTCAACATTGAGTTATTACAACCGTGGATCTTTATAGATACGATCCC | 324 |
| QY | 94 | TrpGluLeuProAlaMetAlaAlaGlyGluLysGluTyrTrpPheTyrValProArg | 113 |
| DB | 325 | TGGGAGCTTCCAGCTTTCGCTTCAATGGAGAGAAGAGTTGTTCTTTTACGTATACGA | 384 |
| QY | 114 | AspArgLysTyrArgAsnGlyAspArgProhenArgValThrSerGlyTyrTrpLys | 133 |
| DB | 385 | GATAGAAAAGTAGAATGGAGATCGACCTAACAGAGTTACAAAGTCTGGGTATTGGAAA | 444 |
| QY | 134 | AlaThrGlyAlaAspArgMetIleArgSerGluThrSerArgProIleGlyLeuLysLys | 153 |
| DB | 445 | GCTACAGTGTGACCGGATGGTTCCAGTGAAGTCTCTGGTTGCATCGCTTGAAGAAA | 504 |
| QY | 154 | ThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIleMetAsn | 173 |
| DB | 505 | ACGCTTGTGTTCTATACAGGGAAGCGCCCAAGGGAAGCAAGCAAGCTGGATCATGACC | 564 |
| QY | 174 | GluTyrArgLeuProHisGluThrGluLysTyrGlnLysAlaGluLeuSerLeuCys | 193 |
| DB | 565 | GAATATAGTGTTCACAACTCGAAACTCGAAATTCAGAAAGCAAGCAACTCCCTTCCTGC | 624 |
| QY | 194 | ArgValTyrLysArgProGlyValGluAspHisProSerValProArgSerLeuSerThr | 213 |
| DB | 625 | CCGTTTCGAAAAGGTGAGCA-----AAATCCGGAAGGAGGTTTGGAGCA | 672 |
| QY | 214 | ArgHisHisAsnHisAsnSerSerThrSerSerArgLeu----- | 226 |
| DB | 673 | AAGGAACATGTTAAGGGAACCGGAACTGATAGCAACACTTAATAAAGATAATGTTGA | 732 |
| QY | 227 | ---AlaLeuArgGlnGlnHisHisSerSerSerSerAsnHisSerAspAsnLeu | 245 |
| DB | 733 | CCTTGTTCATATTGAAATGCAGGAATCAGGTATGTCAGTGGCAGACCTGTCATG | 792 |
| QY | 246 | AsnAsnAsnAsnIleAsnAsnLeuGluLysLeuSerThrGluTyrSerGlyAspGly | 265 |
| DB | 793 | GCATCAACACATTAAGAACACCGTTCAGTCTTGAATTTGGAGATCAAGCA---GGA | 849 |
| QY | 266 | SerThrThrThrThrThrThrAsnSerAsnSerAspValThrIleAlaLeuAlaGln | 285 |

Db 850 AGTACAGATGAATTTGTTAAAGCTCGAGGAATGCCAACAGATATAAAAGTTATGCAAAA 909
 QY 286 Aenlle-----TyrArgProMetProTyrAsp 294
 Db 910 TCAGTAATCTGGGAGGAAAGATACCCAGCTCAAAATGGAGTACCGAGTAACAGATCAGTCG 969
 QY 295 ThrSerAsnAsnThrLeuIleValSerThrArgAsnHis-----GlnAspAspAsp 311
 Db 970 GAACACAGTCAGGAATGCATTACAGCAGACCGGGGAATATTGTAGTAGAGTGAGGCAG 1029
 QY 312 GluThrAlaIleValAspAspLeuGlnArgLeuValAsnTyrGlnIleSerAspGlyAla 331
 Db 1030 CCCATCTCTGTAATGACGAGCTT-----GAACTATCTTTGACGACA 1071
 QY 332 ThrThrLeuMetProGlnThrGlnAlaAlaLeuAlaMet----- 344
 Db 1072 TCCTTAATCTGTCGAGAGAACAGATCTGCCGGCAGTACCAAGATCTCTATAATAAT 1131
 QY 345 AsnMetIleProAlaGlyThrIleProAsnAsnAlaLeuTrpAsp----- 359
 Db 1132 AATAATAATCCATCTCCCTAGTCTCTCAATAATGCTTTAACTGATGGTGGCGTTGGCAA 1191
 QY 360 -----MetTrpAsnProIleValProAspGlyAsn 369
 Db 1192 ACTTTCACTGACAAATATTGGAGTGGAAACCTGTGTGAGAAAGCAATGGCAGT 1242
 RESULT 13
 ID ADA70817 standard; DNA; 1158 BP.
 XX AC ADA70817;
 XX DT 20-NOV-2003 (first entry)
 XX DE Rice gene, SEQ ID 4140.
 XX KW plant; bacterial infection; fungal infection; viral infection; rice;
 XX KW gene; ds.
 XX OS Oryza sativa.
 XX PN WO2003000898-A1.
 XX PD 03-JAN-2003.
 XX PF 22-JUN-2001; 2001WO-IB001105.
 XX PR 22-JUN-2001; 2001WO-IB001105.
 XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
 XX PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
 XX WPI, 2003-175290/17.
 XX Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 XX Claim 6; SEQ ID NO 4140; 899pp; English.
 XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

XX SQ Sequence 1158 BP; 333 A; 295 C; 282 G; 248 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4.68e-42 Length: 1158
 Score: 544.50 Matches: 132
 Percent Similarity: 51.0% Conservative: 65
 Best Local Similarity: 34.2% Mismatches: 134
 Query Match: 26.8% Indels: 55
 Gaps: 11
 US-10-780-703-2 (1-379) x ADA70817 (1-1158)
 QY 33 GlnGluSerHisValGlnAsnGluAspGluAlaAspAspHisAspHisAspMetValMet 52
 Db 4 GAGAGAGAAATGATGTCAACATGACATGACAAAGTCAGAT-----GAGATCTCTCTG 51
 QY 53 ProGlyPheArgPheHisProThrGluGluLeuLeuGluPheTyrLeuArgArgLys 72
 Db CCTGGATTTCAGGTTTCATCTCTACAGATGACAGAGCTGGTTAGCTTCTATCTCAAGAGGAAG 111
 QY 73 ValGluGlyLysArgPheAsnValGluLeuIleThrPheLeuAspLeuTyrArgTyrAsp 92
 Db ATCCAGCAGAAAGCCCTATCTCCATTGAGCTCATCAGGCAGCTTGACATCTCAAGTTTGTAT 171
 QY 93 ProTrpGluLeuProAlaMetAlaAlaIleGlyGluLysGluTyrPheTyrValPro 112
 Db CTTTGGGATCTCCCAAAGCTTGGCAGCAGCGGTGAAAAGAGTGGTACTTCTTACGCCCG 231
 QY 113 ArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrSerGlyTyrTrp 132
 Db AGGACCCGGAATATCGCAACAGCGTGTAGGCGGACCGGTGACAAACAGCTGGGTCTCTGG 291
 QY 133 LysAlaThrGlyAlaAspArgMetIleArgSer---GluThrSerArgProIleGlyLeu 151
 Db AAAGCCACAGAACTGACAGGCCAATCTACTCCACTGAGGGTACCAAGTGCATAGGCCCTC 351
 QY 152 LysLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTrpIle 171
 Db AAGANGTCCCTTGTCTTCTACAGGCGAGAGCTCGCAGAGGAGTCAAGCTGAGTGGATG 411
 QY 172 MetAsnGluTyrArgLeuProHisGluThrGluLysTyrGlnLysAlaGluIle--- 190
 Db ATGCATGAGTTTCAGGCTTCTCTACCGCTCACCCTCATCGCTCCCGAAGACCGATTGAC 471
 QY 191 -----SerLeuCysArgValTyrLysArg----- 198
 Db AAGAACATTCGGCTCAACGACTCTCTGGACATCTCGAGGATCTTCAAGAAAGACCAAGTTCC 531
 QY 199 -----ProGlyValGluAspHis 204
 Db 532 ATGCCACAAGGGCAGCTCTGTACAGCTGGGGAGCTCAATTTGCTGGACCAATTTGAT--- 588
 QY 205 ProSerValProArgSerLeuSerThrArgHisAsnHisAsnSerSerThrSerSer 224
 Db CCAGACATCTTCTACATTCGCAATCTGTGCAAGCTTCACAGTTTGCCTCGAGAGTTCC 648
 QY 225 ArgLeuAlaLeuArg-----GlnGlnGlnHisSerSerSerSerSerSerSerSer 242
 Db TCCTGTCTCATTTGCAAGCTGCGACCAACGGCAGCTCATCAGATCCTAGCAAGTATGTCAT 708
 QY 243 AsnAsnLeuAsnAsnAsnAsnAsnIleAsnAsnLeuGluLysLeuSerThrGluTyrSer 262
 Db 709 CAAGGTAAATAATAATAATACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTACCTAGAT 768
 QY 263 GlyAspGlySerThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 282
 Db 769 GGA-----TCTTCATGCAAGGTCAATAAATTTCAACTGCAGTCAATCTTCGAGAAGTC 819
 QY 283 AlaAsnGlnAsnIleTyrArgProMetProTyrAspThrSerAsnAsnThr----- 299
 Db 820 CAGAACTCCCAAGATCATATGTCCTATTTGAGGCACACACATCGCAGAAAGCCGCTACGCCG 879

QY 300 LeuValSerThrArgAsnHisGlnAspAspGluThrAlaIleValAspLeu 319
 DB 880 CTGCTCTTCACACGCGTTGGCGCTGACAGATGATTTGGTGCAT--- 936
 QY 320 GlnArgLeuValAsnTyrGlnIleSerAspGlyAlaThrThrLeuMetProGlnThrGln 339
 DB 937 -----TCATCTGTAATGCAATGGAGGAGCATCAGCAACAGACCAAGATCCA 987
 QY 340 AlaAlaLeuAlaMetAsnMetIleProAlaGlyThrIleProAsnAsnAlaLeuTyrAsp 359
 DB 988 TCAGCAAGAAAG-----CCTGGCAGTGGCTTCAGCATGAACACAGTGCATGGGAT 1035
 QY 360 -----MetTTPAsnProIleValProAspGlyAsnArgAspHisTyr 373
 DB 1036 GGTGTGGCAAGATCAACTCCGTTGATTTAGGAGCAGATCTTCAGAGATGGAGG 1095
 QY 374 ThrAsnIleProPheLys 379
 DB 1096 AGCAGCATACCTGGGAA 1113
 RESULT 14
 ID ADA70350 standard; DNA; 950 BP.
 AC
 AC ADA70350;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Rice gene, SEQ ID 3673.
 KW Plant; bacterial infection; fungal infection; viral infection; rice;
 KW gene; ds.
 XX
 OS Oryza sativa.
 XX
 PN WO2003000898-A1.
 XX
 PD 03-JAN-2003.
 XX
 PF 22-JUN-2001; 2001WO-IB001105.
 XX
 PR 22-JUN-2001; 2001WO-IB001105.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
 XX
 DR WPI; 2003-175290/17.
 XX
 PT Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 XX
 PS Claim 6; SEQ ID NO 3673; 899pp; English.
 XX
 CC The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC the expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 XX
 SQ Sequence 950 BP; 186 A; 319 C; 312 G; 133 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,316-41 Length: 950
 Score: 536.00 Matches: 107

Percent Similarity: 60.4% Conservative: 44
 Best Local Similarity: 42.8% Mismatches: 81
 Query Match: 26.4% Indels: 18
 DB: 8 Gaps: 5
 US-10-780-703-2 (1-379) x ADA70350 (1-950)
 QY 52 MetProGlyPheArgPheHisProThrGluGluGluLeuLeuGluPheThrLeuArgAsp 71
 DB 52 CTCCCGGGCTTCGGTTTCCACCCACGGAGGAGGAGCTCTCGACTTCCTACCTCTCCCGC 111
 QY 72 LysValGluGlyLysArgPheAsnValGluLeuIleThrPheLeuAspLeuTyrArgTyr 91
 DB 112 GTCGTCTTAGCAAGAGCTCCACTTCAACATCATCGCACCCCTCAACATCTACCGCCAT 171
 QY 92 AspProTTPGluLeuProAlaMetAlaIleGlyGluLysGluTyrPheThrVal 111
 DB 172 GATCCCTGGGACCTCCAGGATGCGAAGATCGGGGAGAGGAGTGGTACTTCTTCGTG 231
 QY 112 ProArgAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyr 131
 DB 232 CCGCGGACAGAGAGCGCGGAGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 291
 QY 132 TrpLysAlaThrGlyAlaAspArgMetIleArgSer-----GluThrSerArgProIle 149
 DB 292 TGGAGGCGACGGGTTCGACAGAGGCGATCCGGAGCTCCGGCGACCCGAGCGGGTGATC 351
 QY 150 GlyLeuLysLysThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSer 169
 DB 352 GGGCTCAAGAGACACGCTCGTCTTACCAAGGTCGCGCGCGCGCGCGCGCGCGCGCGAC 411
 QY 170 TrpIleMetAsnGluTyrArgLeuProHisHisGluThrGluLysTyr----- 185
 DB 412 TGGGTGATGAACAGATACCGCTCCCGACTACGGCGCGCGCGCGCGCGCGCGCGCGCT 471
 QY 186 GlnLysAlaGluIleSerLeuCysArgValTyrLysArg-----ProGlyVal 201
 DB 472 CCCAAGGAGGACATGGTGTCTCTGCAAGATATATACCGAAGCGCGCGCGCGCGCGCTG 531
 QY 202 GluAspHisProSerValProArgSerLeuSerThrArgHisHisAsnHisAsnSerSer 221
 DB 532 GAGCAGAGAGCGCTCGGCAATGGAAGAGATGCAG---AGAGGATCAAGTCACGAGACTAC 588
 QY 222 ThrSerSerArgLeuAlaLeu-----ArgGlnGlnGlnHis 234
 DB 589 ACGGCGACGAGAGCATCTCTCTCCACGAGCGCTCGCTCCACGCGCGCGCGCGCGCGCTTC 648
 QY 235 SerSerSerSerAsnHisSerAspAsnAsnLeuAsnAsnAsnAsnAsnLeu 254
 DB 649 TCGTCGGAGCGCTCCACGACGAGCGGCTTCTGTATCCAGTCTCATCTCGTCGCGCGG 708
 QY 255 GluLysLeuSerThrGluTyrSerGlyAspGlySerThrThrThrThrThrThrAsnSer 274
 DB 709 CGTCCGCGGAGCAGCAGCAAGAACCGCGCGCTGGCGCGCGCGCGCGCGCGCGCGCGAA 768
 QY 275 AsnSerAspValThrIleAlaLeuAlaLeu 284
 DB 769 GAAGCAGACGTCACGTCACGTCACGTCGCGCTG 798
 RESULT 15
 ID ADA70350 standard; cDNA; 1644 BP.
 AC ADA70350;
 XX
 DT 21-APR-2005 (first entry)
 XX
 DE Plant full length insert polynucleotide seqid 20520.
 KW plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;

growth rate; cell cycle pathway; disease resistance;
 galactomannan production; lignin production; plant growth regulator;
 yield; plant growth; plant development; seed oil; protein yield;
 protein content; gene; ss.
 Unidentified.
 US2004034888-A1.
 19-FEB-2004.
 28-APR-2003; 2003US-00425114.
 06-MAY-1999; 99US-00304517.
 05-NOV-2001; 2001US-00985678.
 (LIU//) LIU J.
 (ZHOU//) ZHOU Y.
 (KOVA//) KOVALIC D K.
 (SCRE//) SCREEN S E.
 (TAB//) TABASKA J E.
 (CAO//) CAO Y.
 Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 MPI; 2004-180133/17.
 New recombinant DNA construct, useful for improving plant tolerance to
 cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 pests, for conferring increased resistance to plant disease, or for
 improving yield.
 Claim 1; SEQ ID NO 20520; 15pp; English.
 The invention describes a recombinant DNA construct comprising a
 polynucleotide consisting of a sequence encoding an amino acid sequence
 available in electronic form from the US patent office at
 ftp://seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
 of the invention are also useful in physical arrays of molecules and as
 plant breeding markers. The recombinant DNA construct is useful for
 improving plant tolerance to cold, heat, drought, herbicides, extreme
 osmotic conditions, pathogens or pests, for manipulating growth rate in
 plant cells by modification of the cell cycle pathway, for conferring
 increased resistance to plant disease, for producing galactomannan,
 lignin or plant growth regulators, for increasing the rate of homologous
 recombination in plants, for improving yield by modification of
 photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake,
 or by providing improved plant growth and development under at least one
 stress condition or for modifying seed oil or protein yield and/or
 content. This sequence represents a plant full length insert
 polynucleotide that can be used in the recombinant DNA construct of the
 invention.
 Alignment Scores:
 Pred. No.: 6,07e-41 Length: 1644
 Score: 535.00 Matches: 129
 Percent Similarity: 50.9% Conservative: 73
 Best Local Similarity: 32.5% Mismatches: 130
 Query Match: 26.3% Indels: 66
 DB: 13 Gaps: 12
 US-10-780-703-2 (1-379) x ADX45780 (1-1644)
 14 SerAsnGlnValAsnAsnGluLysGlyIleGluAspAsnAspHis-ArgGlyGlyG1 33
 90 AACACACATAGTGAACACACAC-----CTTGGTAATATATACCATGGAGACCGCA 140
 33 nGluSerHisValGlnAsnGluAspGluAlaAspAspHisAspMetValMetPr 53
 141 AGACTCTCCAGGAGTTCACGACCAACCGCAC-----CTGCCACC 182

53 oGlyPheArgPheHisProThrGluGluGluLeuIleGluPheTyrLeuArgAlaGlyVa 73
 183 CGGGTTTCGGTTCCACCCACAGACGAGGAGCTGGTTTCATTACCTCCTCAAGAGAAAGT 242
 73 lGluGlyLysArgPheAsnValGluLeuIleThrPheLeuAspLeuTyrArgTyrAspPr 93
 243 AGATTCTGTTCT 302
 93 oTTPGluLeuProAlaMetAlaIleGlyGlyGluLysGluLysGluLysGluLysGlu 113
 303 ATGGGAATTACCGCCAAAGCATCGTTTGGGGCGGAGAGGGTACTTTTTCAGCCCGAG 362
 113 gAspArgLysTyrArgAsnGlyAspArgProAsnArgValThrThrSerGlyTyrTrpLy 133
 363 AGAAGGAAGTACCGCAATGGCGGAGGCCCAACAGGGCTCAACTTCAGGGTATTGGAA 422
 133 sAlaThrGlyAlaAspArgMetIleArgSerGluThrThrSerArgProIleGlyLeuLysLy 153
 423 AGCCACGGGACAGATAAGCCTATATGTAGTGGAACTCAAAAG---GTTGGGGTGAAGAA 479
 153 sThrLeuValPheTyrSerGlyLysAlaProLysGlyThrArgThrSerTyrIleMetAs 173
 480 ATCTTTGGTTTCTATGGAGGAGGCCCAACCAAGGGGTAAACCCGATTGGATCGCA 539
 173 nGluTyrArgLeuProHis-----Hi 180
 540 TGAGTACGGTGTACTGAAACCAAGCCCAATAGGCCTCTCGGTGTGACTTGGGCCA 599
 180 sGluThrGluLysTyrGlnLysAlaGluIleSerLeuCysArgValTyrLysArgProGl 200
 600 CAAGAAAACCTCCCTTAAGGGCTGGATGATTGTGGTGTGTGGCGAATCTCAAGAAAGGCAA 659
 200 yValGlu---AspHisPro-----SerValProAr 209
 660 CACACAAAGGTCACATGAGAGGAGCATTCATTGATGACATGATGAGAGAGTACCTCC 719
 209 gSer-----LeuSerThrArgHisHisAsnHisAsnSerSerThrSerSe 224
 720 TTCATCAACGTTGGGGCCACATGATGCGAGGTTTCACCTTTCAAGGATGTCACGAGCTA 779
 224 rArgLeuAlaLeuArgGlnGlnHisHisSerSerSer-----SerAsnHi 240
 780 CAGTGGTGCATTTGTGGAAATGACAGAAACACACATAGAGGGGTGTTATAGGCAATGG 839
 240 sSerAspAsnAsnLeuAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsn 260
 840 TAGTGTGAATGGGATGAGCAATCTAATAATATATACAGTGTCTATCATCTCATCATCA 899
 260 uTyrSerGlyAspGlySer-----ThrThrThrThrThrTh 272
 900 GTTTGGCACCTCAAACTCCAAAGGCGAGAGCTTCTCTTTTGTCTTCCCTTCAAAACACATTC 959
 272 rAsnSerAsnSerAspValThrIleAlaLeuAlaAsnGlnAsnIle-----TyrArgPr 290
 960 TAATTGGCGCTCCAAAGAGACACTCTCATCTACTTATTTGGAAATGTTGATGATCACAACA 1019
 290 oMetProTyrAspThrSerAsnAsnThrLeuIleValSerThrArgAsnHisGlnAspAs 310
 1020 CTTCAATTGGACAGATTAATGGGAATGTAGTGTGTTTGGAGACAGATCAAGAGAAATAATAA 1079
 310 pAspGluThrAlaIleValAspAspLeuGlnArgLeuValAsnTyrGlnIleSerAspGl 330
 1080 TAATAATAGGTACTTCTGGCTCTCTTGTGCTCTCTGCTTACTCTGCTTAAACCAG----- 1125
 330 yAlaThrThrLeuMetProGlnThrGlnAlaAlaLeuAlaMetAsnMetIleProAlaGl 350
 1126 -----CTTCCACAAACACCTTCATCTGCCCAATAGATGG-GTTACTTCGGA 1171
 350 yThrIleProAsnAsnAlaLeuTyrAspMetTyrAsnProIleValPro 366
 1172 CACCGTATCAATACAAAGGGACCAATTGGTATGG-----TTAATTCGA 1214

Search completed: April 6, 2006, 11:03:49
Job time : 644 secs
